



A University of Sussex DPhil thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

Multiparametric Interfaces For Fine-Grained Control of Digital Music

Chris Kiefer

Submitted for the degree of D.Phil.

University of Sussex

September, 2012

DECLARATION

I hereby declare that this thesis has not been submitted, either in the same or different form, to this or any other university for a degree.

Signature:

Supervisors: Dr. Nick Collins, Dr. Geraldine Fitzpatrick (between 2007 and 2009) and Dr. Graham McAllister (between 2009 and 2011)

Examiners: Dr. Nick Bryan-Kinns and Dr. Judith Good

for Tinne, Astrid and Frida

ACKNOWLEDGEMENTS

I have massive, massive gratitude and bear a huge debt to Tinne and our daughter Astrid. Tinne has given me so much support through all of this thesis, and in taking up slack while I've been so busy, has probably worked as hard on it as I have. Along with this, they've provided enough fun and madness at home to keep me relatively sane throughout. Thank you also to all my family and friends.

I am immensely grateful to my supervisors Geraldine Fitzpatrick, Graham McAllister and especially my main supervisor Nick Collins, who have been encouraging, supportive and inspiring to work with. I've been lucky enough to share my time at Sussex with some great people, thanks to everyone in the Interact Lab, and to my fellow Music Informatics researchers: Thor Magnusson, Matthew Yee-King and also Anna Jordanous, with whom I've shared the (mostly) ups and downs of the DPhil journey.

I'd like to thank Jon Bird for organising an Arduino workshop in 2007 that put electronics into a new light. Inman Harvey allowed me to sit in on the inspiring A-Life course and also gave me some valuable algorithmic advice. Thank you to Cathy Treadaway for some great discussion about hands, art and creativity. Also, huge thanks to Prathaban Ramalingam for giving me profound advice and moral support.

Most of my experiences in electronic music that lead up to the start of this thesis were shared with Liz Green and Andrew Stock, I will always be so grateful to them for this. Thanks also to everyone at Wasp Factory, Punish and Cryonica.

Much of this thesis has been based on free software, so I must acknowledge the programmers behind SuperCollider, OpenFrameworks, LiveOSC and Freemind for the generous amounts of time they have put into these projects. Along with this, thank you to everyone on the sc-users list and the OpenFrameworks forum who have spared the time to give me technical advice.

The University of Sussex and the EPSRC funded my DPhil; Sussex has been a lovely environment to work in and provided great training and support for a confused fledgling PhD student.

Massive thanks to the many participants who spared the time and energy to test out some new music controllers, especially to those who took part in the field studies.

Multiparametric Interfaces For Fine-Grained Control of Digital Music

Chris Kiefer

Summary

Digital technology provides a very powerful medium for musical creativity, and the way in which we interface and interact with computers has a huge bearing on our ability to realise our artistic aims. The standard input devices available for the control of digital music tools tend to afford a low quality of embodied control; they fail to realise our innate expressiveness and dexterity of motion. This thesis looks at ways of capturing more detailed and subtle motion for the control of computer music tools; it examines how this motion can be used to control music software, and evaluates musicians' experience of using these systems.

Two new musical controllers were created, based on a multiparametric paradigm where multiple, continuous, concurrent motion data streams are mapped to the control of musical parameters. The first controller, Phalanger, is a markerless video tracking system that enables the use of hand and finger motion for musical control. EchoFoam, the second system, is a malleable controller, operated through the manipulation of conductive foam. Both systems use machine learning techniques at the core of their functionality. These controllers are front ends to RECZ, a high-level mapping tool for multiparametric data streams.

The development of these systems and the evaluation of musicians' experience of their use constructs a detailed picture of multiparametric musical control. This work contributes to the developing intersection between the fields of computer music and human-computer interaction. The principal contributions are the two new musical controllers, and a set of guidelines for the design and use of multiparametric interfaces for the control of digital music. This work also acts as a case study of the application of HCI user experience evaluation methodology to musical interfaces.

The results highlight important themes concerning multiparametric musical control. These include the use of metaphor and imagery, choreography and language creation, individual differences and uncontrol. They highlight how this style of interface can fit into the creative process, and advocate a pluralistic approach to the control of digital music tools where different input devices fit different creative scenarios.

Submitted for the degree of D.Phil.

University of Sussex

September, 2012

CONTENTS

I	Introduction	I
1.1	Interacting With Digital Music Tools	2
1.2	Research Context	3
1.3	Research Questions and Themes	3
1.4	Contributions	4
1.5	Relevance	4
1.6	Structure	5
1.7	Related Publications	6
1.8	Working Definitions	6
1.8.1	Musicality	6
1.8.2	Intuitive Interaction	7
2	Musicality and Digital Music Tools	9
2.1	Introduction	10
2.2	Interaction in Composition and Performance	10
2.2.1	Where Composition Meets Performance	11
2.2.2	Composer Studies	11
2.2.3	Composing Instruments	12
2.2.4	Musical Activity	13
2.3	Modes Of Interaction	13
2.4	Computer Music Instrumentation	14
2.4.1	The Analogue Studio as an Instrument	14
2.4.2	The Digital Music Studio as an Instrument	15
2.4.3	Disconnection	17
2.4.4	Reconnection	20
2.4.4.1	Recent Directions in Academic Research	20
2.4.4.2	Multi-touch and Mobile Technology	21
2.4.4.3	DIY Music Culture	21
2.5	The Body In Interaction	21
2.5.1	Embodiment and Human-Computer Interaction	22
2.5.2	Embodiment and Computer Music	24
2.6	A Holistic Perspective of the Computer Music Instrument	25
2.7	Problematic Areas In Musical Interaction	26

2.7.1	Degree of Embodiment	26
2.7.2	Conceptual Complexity	27
2.7.3	Analysing the Chart	27
2.8	Conclusion	28
3	Design Aims	30
3.1	Introduction	31
3.2	Designing for Creativity	31
3.2.1	Design Principles for Creativity Support Tools	31
3.2.2	Flow	32
3.3	Desirable Qualities in Musical Instruments	34
3.4	Capturing Embodied Interaction	36
3.4.1	Degrees of Freedom of Motion	37
3.4.2	Multiparametric Control	38
3.4.3	A Convergent Model of Interaction	39
3.5	A Design Space	40
3.6	Emergence and Real-world Constraints	41
3.7	Summary	41
4	HCI Methodology for Evaluating Musical Interfaces	43
4.1	Introduction	44
4.1.1	The Challenges of Evaluating Musical Controllers	44
4.2	Evaluating the Wiimote	45
4.2.1	The Wiimote	46
4.2.1.1	The Accelerometer	46
4.2.2	Methodology	47
4.2.3	The Study	47
4.2.3.1	Metaphors and Musical Tasks	47
4.2.3.2	Participants	49
4.2.3.3	Method and Implementation	49
4.2.3.4	Implementation	49
4.2.4	Results	49
4.2.4.1	Quantitative Results	49
4.2.4.2	Interview Results	50
4.2.5	Discussion	52
4.2.6	Summary	52
4.2.7	Reflecting on the Wiimote Study	52
4.3	Evaluation Methodology	53
4.3.1	From Usability to User Experience	53
4.3.2	Physiological Measurements	54
4.3.3	Real-world Musical Interaction - Evaluation In The Wild	54
4.3.3.1	Data Collection	56
4.3.4	Combining Research Methods	56

4.3.5	Data Analysis	57
4.3.5.1	Grounded Theory	57
4.3.5.2	Grounded Theory Process	58
4.4	Summary	59
5	Phalanger: A Hand Tracking Interface	60
5.1	About Phalanger	61
5.1.1	Related Work	61
5.1.2	Implementation	62
5.1.2.1	Background Segmentation	62
5.1.2.2	Frame Analysis	63
5.1.2.3	Hand Position Detection	64
5.1.2.4	In Use	65
5.1.3	Evaluation	65
5.2	Formative Evaluation	65
5.2.1	Results	67
5.2.2	Discussion	68
5.2.3	Summary	68
5.3	A Field Evaluation of Phalanger	69
5.3.1	Implementation Changes	69
5.3.1.1	Ergonomic Improvements	69
5.3.1.2	Performance Improvements	70
5.3.2	Field Study Participants	70
5.3.3	Software	70
5.3.3.1	Common Features	70
5.3.3.2	The OSC Version	71
5.3.3.3	The MIDI Version	71
5.3.3.4	Other Software	71
5.3.4	The Study	74
5.3.5	Data Analysis	75
5.3.6	Results	75
5.3.6.1	Concerning Phalanger	75
5.3.6.2	The Physical	77
5.3.6.3	Interaction	78
5.3.6.4	Making Music	78
5.3.7	Discussion	79
5.3.7.1	Scope and Limitations of the Study	79
5.3.7.2	Discussion of Results	79
5.4	Summary	80
6	Echofoam: A Malleable Controller	81
6.1	Introduction	82
6.2	EchoFoam	83

6.2.1	Sensor Construction	83
6.2.2	Mapping with Echo State Networks	84
6.2.3	Software Implementation	86
6.2.4	ESN Training	86
6.3	Formative Evaluation	87
6.3.1	Scenarios	88
6.3.2	Participants	89
6.3.3	Method	89
6.3.4	Results	89
6.3.5	Discussion	93
6.4	Design Improvements	94
6.4.1	Packaging	94
6.4.2	Increasing controller responsiveness	95
6.4.3	Foam Sensitivity Degradation	97
6.5	Field Evaluation	97
6.5.1	Software	99
6.5.1.1	Implementation	100
6.5.1.1.1	LiveOSC	100
6.5.1.1.2	Tools	101
6.5.1.1.3	Echo State Network Mapping	101
6.5.1.2	Features	102
6.5.1.2.1	Parameter Selection	102
6.5.1.2.2	Search Features	102
6.5.1.2.3	User Interface Features	102
6.5.2	Study Design	102
6.5.2.1	Participants	102
6.5.2.2	Method	105
6.5.3	Data	105
6.5.4	Results	105
6.5.4.1	The Controller	105
6.5.4.2	Aspects of Control	108
6.5.4.3	Software Use	110
6.5.4.4	Mapping	111
6.5.4.5	The System in Use	114
6.5.5	Discussion	117
6.5.5.1	Methodology	117
6.5.5.2	Analysing the Results	118
6.5.6	Conclusions	119
6.6	Summary	119

7	Multiparametric Exploration of Timbre Spaces	121
7.1	Introduction	122
7.1.1	Research Questions	122
7.1.2	Unpredictable Control	123
7.2	Related Work	123
7.3	System Architecture	124
7.4	Exploring Timbre Navigation Strategies	124
7.4.1	Scenario 1: A Four Dimensional Timbre Space	124
7.4.2	Scenario 2: A Ten Dimensional Timbre Space	127
7.4.3	Scenario 3: A Forty Dimensional Timbre Space, Navigated With 10 Control Streams	127
7.5	Initial Reactions	127
7.5.1	Discussion	128
7.6	Evaluation	129
7.6.1	Participants	129
7.6.2	Scenario	129
7.6.3	Method	130
7.6.4	Results	130
7.6.4.1	About the System	130
7.6.4.2	Control	132
7.6.4.3	About RECZ	133
7.6.4.4	Player Experience	134
7.6.4.5	Methodology	136
7.7	Discussion	136
7.8	Conclusions	137
8	Discussion	138
8.1	Looking Back	139
8.2	Review of Evaluation Methodology	139
8.2.1	Formative Evaluation Methods	139
8.2.2	Field Evaluation Methods	141
8.2.3	Combined Methods	142
8.2.4	Evaluation Concepts	142
8.3	Combined Results	143
8.3.1	Phalanger	143
8.3.1.1	Convergent Interaction Model	143
8.3.1.2	Open Design	145
8.3.1.3	Learning Curve	145
8.3.1.4	Continuous Control	146
8.3.1.5	Intuitive Control	146
8.3.1.6	Fine-grained Control	146
8.3.1.7	Uncontrol	146
8.3.1.8	Summary	146

8.3.2	EchoFoam	147
8.3.2.1	Convergent Interaction Model	147
8.3.2.2	Open Design	147
8.3.2.3	Learning Curve	147
8.3.2.4	Continuous Control	148
8.3.2.5	Intuitive Control	148
8.3.2.6	Fine-grained Control	148
8.3.2.7	Uncontrol	148
8.3.2.8	Summary	148
8.4	Emerging Themes	149
8.4.1	Metaphor, Imagery and Sound-Motion Correspondence	149
8.4.2	Choreography and Language Creation	150
8.4.3	Cognitive Styles and Multiparametric Control	151
8.4.4	Uncontrol	153
8.4.4.1	A Mini-Taxonomy of Uncontrol	154
8.4.4.2	Damping Nonlinear Behaviour	156
8.4.4.3	Summarising Uncontrol	156
8.4.5	The Creative Process	157
8.5	Summary	158
9	Conclusions	159
9.1	Towards a Model of Multiparametric Interaction	160
9.2	Future Work	162
9.2.1	Phalanger	162
9.2.2	EchoFoam	163
9.2.3	A Generic Multiparametric Template	163
9.2.4	Mapping Techniques	163
9.2.5	Embodied Control and Multiparametric Controllers	163
9.3	Summary	164
A	Building an EchoFoam Controller	166
A.1	Introduction	167
A.2	Materials and Tools	167
A.3	Instructions	167
A.3.1	Making the Controller	167
A.3.2	Electronics	169
A.3.3	Testing	169
B	EchoFoam Live Link Manual	171
B.1	About	172
B.2	Getting Started	173
B.2.1	Hardware Setup	173
B.2.2	Software Setup	173

B.3	The EchoFoam Controller	173
B.4	In Use	174
B.4.1	Cautionary Advice	174
B.4.2	Scope / Limitations	174
B.4.3	Overview	174
B.4.4	User Interface	175
B.4.5	Choosing Parameters	177
B.4.5.1	Choosing Single Parameters	177
B.4.5.2	Populating with Randomly Chosen Parameters	177
B.4.6	Controlling the Search Space	177
B.4.7	Key Commands	178
B.4.8	Adjusting controller sensitivity	178
B.4.9	Using the Arduino for other projects	179
C	Collected Concepts from Interview Analyses	180
C.1	Collected Concepts from all Interview Results	181

LIST OF FIGURES

1.1	Thesis Structure	5
2.1	King Tubby	14
2.2	Orbital Live, 2010	16
2.3	Siemens Studio for Electronic Music, ca. 1956	18
2.4	Roland System 100	18
2.5	An Akai S1000 Sampler	19
2.6	Steinberg Cubase SX, Digital Audio Workstation software	19
2.7	Embodiment vs Conceptual Complexity	27
3.1	Design Scope	31
3.2	Flow: Challenge vs Skill	33
3.3	A Convergent Model of Interaction	39
3.4	Design Space	40
4.1	The Nintendo Wiimote	46
4.2	The HandSonic and the Wiimote Used in the Study	48
5.1	An Overview of the Phalanger System	62
5.2	Frame Analysis in Phalanger	63
5.3	Evaluation Scenario 1: A Tetris Style Sound Game	66
5.4	Evaluation Scenario 2: Hand Shape Mapped to Granular Synthesis	66
5.5	Evaluation Scenario 3: A Virtual Mixer	67
5.6	Phalanger In Use With A Desk Facing Camera	69
5.7	Phalanger MIDI Configuration	73
5.8	Macam	74
5.9	Phalanger Field Evaluation Interview Analysis: Concepts and Categories	76
6.1	An EchoFoam Controller	84
6.2	An Example ESN	85
6.3	System Overview	86
6.4	Echo State Network Mappings	87
6.5	Kenton Control Freak MIDI Controller	88
6.6	The Controller In Use	90
6.7	An EchoFoam Controller, Version 2	94
6.8	EchoFoam Arduino Shield	95

6.9	Multiplexing in the EchoFoam Controller	96
6.10	Enamelled Copper Wire Tied into a Section of the Controller	97
6.11	The Foam Sections of the Controller, Before Being Glued Together	98
6.12	Lines of Measurement in an EchoFoam Cube	98
6.13	EchoFoam Live Link Screenshot 1	99
6.14	EchoFoam Live Link Data Flow	100
6.15	EchoFoam Live Link Screenshot 2: Showing Data Streams	103
6.16	EchoFoam Field Evaluation Interview Analysis: Main Categories	106
6.17	EchoFoam Field Evaluation Interview Analysis: The Controller	106
6.18	EchoFoam Field Evaluation Interview Analysis: Aspects of Control	108
6.19	EchoFoam Field Evaluation Interview Analysis: Software Use	111
6.20	EchoFoam Field Evaluation Interview Analysis: Mapping	112
6.21	EchoFoam Field Evaluation Interview Analysis: The System In Use	114
7.1	System Architecture	125
7.2	Exploring A Timbre Space	126
7.3	Results: Concepts and Categories	131
8.1	Evaluation Concepts	144
8.2	Uncontrol	155
A.1	Connectors	168
A.2	EchoFoam Stripboard Layout	169
A.3	EchoFoam Circuit Schematic	170
A.4	The Finished Circuit Board	170
B.1	EchoFoam Live Link Screenshot	172
B.2	EchoFoam Live Link User Interface	175

LIST OF TABLES

4.1	Controller Preferences	50
5.1	Outputs of the Frame Analysis Process	64
5.2	Phalanger OSC messages	72
5.3	Phalanger MIDI output	73
6.1	Constrained Parameter Ranges for Scenario Three	89
6.2	Echo State Network Configuration	101

CHAPTER I

INTRODUCTION

‘In computer music it has been relatively difficult to provide a friendly and attractive environment for the composer, not because it is so technical, but because we have not spent a lot of time working on the human interface. We have been busy just making sound. This is starting to change.’

James. A. Moorer (Roads, 1982)

1.1 Interacting With Digital Music Tools

The computer can act as an extremely powerful tool for creativity and expression. It can act as a mediator, transforming ideas into tangible forms. It can also provide worlds to explore, within which to inspire and mould ideas. Conversely, the computer can also frustrate, block, divert and misrepresent creative intentions. Ideally, technology should help to realise ideas without subverting them, and do this in a way which encourages engagement and even provides inspiration. This dichotomy is the essence of this thesis; computers supply us with a vast amount of creative potential, but how do we make the most of this valuable resource? The problem now rarely concerns the amount of computing power we have available, but concerns how we use it. The core of this issue is interaction.

Jensenius tells us that *music is movement*, describing how motion is a fundamental to the performance and perception of music (Jensenius, 2007). If we narrow this down to the creation of music, then perhaps it can be said that *music creation is interaction*. Music is produced from our engagement with tools or instruments to create sound, whether this is with a drum, a cello or a digital audio workstation. The instrument is not simply a one-way conduit for musical expression, but an element in a complex, emergent, multi-sensory conversation between player and instrument (Armstrong, 2006).

Acoustic instruments have evolved over centuries or in some cases millennia, and provide us with what many would consider to be an ideal, ground truth model of musical interaction. The player is intimately engaged with the instrument, they experience an embodied connection and sound output bears a direct correlation to physical input. New technology makes this model much more complex; creating music electronically is a relatively new concept, and distinct from the acoustic world because it allows output to be divorced from physical effort. This is an important issue, because interaction now no longer has to follow the rules of the natural world, and the instrument designer has a far greater bearing on the form of conversation that takes place between player and instrument. Digital technology emphasises this issue even further. Computer music systems can exhibit massive complexity and multi-layered abstractions, together with many options for control. The designer now has a huge freedom to create the manner in which the player will converse with the instrument, to design interaction. This gives them the potential to create wonderful, inspiring instruments, but also flawed instruments as well. Understanding the complex nature of how musicians interact with computers is the key to realising their creative potential.

Of the many issues surrounding interaction design for computer music tools, embodiment lies at the core. This is also true for the wider field of human-computer interaction (HCI), with Dourish proposing that embodiment should be the central concern for the field (Dourish, 2001). This perspective focuses on how humans are coupled with the environment, and in turn focuses on how humans should be coupled to technology. It creates an emphasis on the realisation of our innate perceptual-motor skills in interaction.

In improving the quality of embodied interaction with computers, one can look to more complex and detailed techniques for capturing motion as input, relative to the standard devices available for computer musicians. Baseline devices include the mouse and keyboard, and

conventional MIDI controllers, mainly based on scientific-style instrumentation. These devices typically afford limited interaction bandwidth, giving the user less potential to convey the subtlety and detail of natural body motion. The multiparametric approach to controller design attempts to address this issue. Multiparametric controllers use multiple, continuous and concurrent control streams as input, realising our capacity for dexterity and expressiveness of motion, and rich user actions (Djajadiningrat, Matthews, and Stienstra, 2007).

This brings us to the principal focus of this thesis, which is the use of multiparametric control for digital music tools. Hunt and Kirk began research into this field, and found the multiparametric approach to be very promising for musical control (Hunt and Kirk, 2000). Since this study, little research has been explicitly conducted in this area. This thesis looks at how the multiparametric approach can be applied to the control of digital music tools in order to improve musical experience. It looks at multiparametric controllers from a range of angles: building hardware, software mapping techniques, and most importantly, evaluating musicians' experiences of using these controllers. Through these different lenses, I attempt to contribute to a wider picture of multiparametric control that highlights its potential strengths, exposes its potential pitfalls, and informs the computer musician in how to benefit from these techniques in their own work.

1.2 Research Context

This thesis lies on the intersection between two highly multidisciplinary fields: human-computer interaction and computer music, such that it draws on a wide range of academic disciplines, but also from some very specific threads within these areas. From within HCI, it draws on research concerning the philosophy of interaction, recent developments in tangible and physical computing, the application of machine learning to interaction design, and user experience evaluation. Computer music informs elements concerning mapping techniques, sound synthesis and the nature of musical creativity. For simplicity, this research area can be defined as musician-computer interaction; it concerns the nature of interaction between musicians and machines, and the design, implementation and evaluation of technology that facilitates this.

1.3 Research Questions and Themes

The overarching research theme is to investigate how to enhance the experience of interacting with digital music tools. This thesis takes the view that improving the quality of embodied interaction is one possible answer, and that multiparametric interfaces offer a good way of capturing detailed motion for input. With this in mind, the key themes and questions are:

1. Multiparametric control of digital music tools. Multiparametric control is a relatively unexplored area in computer music in the context of controlling digital music tools. How do computer musicians feel about this novel style of interaction? Are more complex, embodied interfaces suitable for controlling music software, and if so, which software processes benefit most from multiparametric control?

2. Evaluation of new musical controllers. There is no standardised methodology for user experience evaluation for digital music interfaces. They present unique challenges that preclude the use of many conventional HCI methods, and expose deeper questions about the design and evaluation cycle when designing for creative use.¹ How can current HCI methodology be adapted for this task?

These questions will be explored through the design and evaluation of two new multiparametric controllers, with final conclusions in chapter 9.

1.4 Contributions

The principle contributions of this thesis are as follows:

- Two new multiparametric musical interfaces; one controlled by video tracked hand motion and the other controlled by manipulation of malleable foam.
- Guidelines for the design and use of multiparametric interfaces for the control of digital music tools, based on in-depth analysis of real-world use by computer musicians.
- A case study of methodology for the evaluation of player experience with new computer music controllers.
- The use of Echo State Networks (see section 6.2.2) as mapping engines for input devices.

1.5 Relevance

Within the span of time I have worked on this thesis, affordable interface technology has taken some significant leaps. At the beginning, musicians were experimenting with the Wiimote and cheap accelerometers. Now, multitouch interfaces are in common use, and the Kinect 3D camera has very recently been released. Coupled with this, we have an exponentially growing scene of hackers and experimentalists who are building their own custom musical controllers, fuelled by the availability of cheap interfaces such as the Arduino and Maple, and growing communities on the web. It would be fruitless to try and keep up with specific technology; what is more important is how we use technology to create interesting and novel musical systems. We need to understand generic principles behind applying new sensors, and we need to understand musicians' behaviour, preferences and creative process. The work in this thesis aims to contribute to a wider picture of musical interaction, and it does this through testing technology with experienced musicians in everyday scenarios. It offers some practical techniques for designers, in the fields of hardware creation, mapping and evaluation. It is my hope that this work will be a useful resource for creators of new controllers who are interested in capturing natural and detailed motion for musical control.

¹Buxton (1997) points out the fundamental difficulty of designing tools for creativity: '*... I ended up learning a lot about design for the artist, since the system that we developed was used by, and influenced by, musicians from around the world. I also discovered that in the grand scheme of things, there are three levels of design: standard spec., military spec., and artist spec. Most significantly, I learned that the third was the hardest (and most important), but if you could nail it, then everything else was easy.*'

1.6 Structure

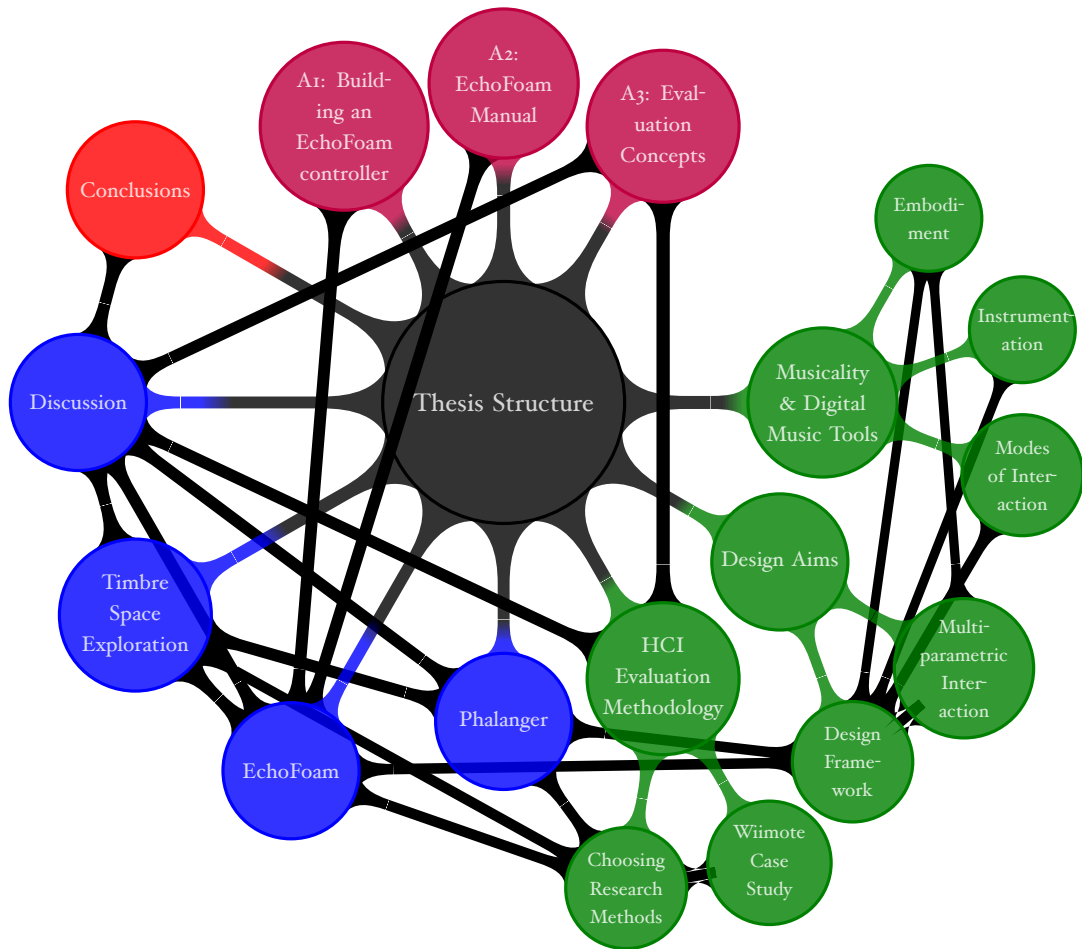


Figure 1.1: Thesis Structure

Chapter 2, examines the nature of interaction between musicians and digital music tools, the importance of embodiment, and sets out the motivations behind this thesis and specific areas of focus. In chapter 3, desirable qualities of musical instruments are explored and a design scope is proposed for new controllers. Chapter 4 explores the use of HCI evaluation methods for musical controllers. A case study of the evaluation of the Wiimote as a musical controller is presented, and evaluation methods for the studies in this thesis are proposed. The next three chapters present new multiparametric controllers and accompanying software systems. Chapter 5 presents *Phalanger*, a video based hand tracking system for musical control. Chapter 6 presents the design and evaluation of the *EchoFoam* system, a malleable foam controller. In chapter 7, the development and evaluation of the *RECZ* software system is explored. *RECZ* was designed as a multiparametric timbrespace exploration tool, which could use both *Phalanger* and *EchoFoam* as input devices. The combined results from the evaluation studies are analysed in chapter 8, and the conclusions follow in chapter 9.

1.7 Related Publications

Some parts of this thesis contain material that has already appeared in conference publications:

Chapter 4 contains material from

- *HCI Methodology for Evaluating Musical Controllers: A Case Study*. In Proceedings of New Interfaces for Musical Expression, Genoa, 2008 (Kiefer, Collins, and Fitzpatrick, 2008a).
- *Evaluating the Wiimote as a Musical Controller*. In Proceedings of the International Computer Music Conference, Belfast, 2008 (Kiefer, Collins, and Fitzpatrick, 2008b).

The formative evaluation of Phalanger, in chapter 5, is described in *Phalanger: Controlling Music Software With Hand Movement Using a Computer Vision and Machine Learning Approach*. In New Interfaces For Musical Expression, Pittsburgh, 2009 (Kiefer, Collins, and Fitzpatrick, 2009).

Formative evaluation of the EchoFoam system, from chapter 6, is described in *A Malleable Interface for Sonic Exploration*. In Proceedings of New Interfaces for Musical Expression, Sydney, 2010 (Kiefer, 2010c).

The development of the timbre space navigation system from chapter 7 is described in *Exploring Timbre Spaces with Two Multiparametric Controllers*. In Proceedings of Sound and Music Computing, Barcelona, 2010 (Kiefer, 2010b).

An overview of research progress was given in *Input devices and mapping techniques for the intuitive control of composition and editing for digital music*. In Proceedings of Tangible, Embedded, and Embodied Interaction, Cambridge, Massachusetts, 2010 (Kiefer, 2010a).

1.8 Working Definitions

The concepts of intuitiveness and musicality crop up frequently in this thesis. These concepts suffer from absent, vague or conflicting definitions in the context of interaction and computer music. To avoid ambiguity, I would like to start by explaining what I understand these terms to mean, and how I arrived at this thinking. I don't claim to eclipse or improve other definitions which already exist, but instead propose working definitions that will be useful in the context of this research. Regardless, when these concepts are mentioned in this text, they refer to these definitions.

1.8.1 MUSICALITY

Starting with a dictionary definition, musicality is curiously undefined in the Oxford Dictionary of Music, yet it is present in the more general Oxford English Dictionary (Online, 2010), as '*The quality or character of being musical; accomplishment or aptitude in music; musical sensibility*'. Perret (2004) considers musicality to be the '*sum total of felt qualitative aspects in musical expression ... Musicality may be considered as being the capacity to draw on a wide spectrum of parameters and*

qualitative aspects in musical expression, which can move and nourish the listener on a deep level. Humans are considered to have innate musicality; this is present across all cultures, and aspects emerge early in life with minimal exposure to music (McDermott, 2008; Trehub, 2003). There are two sides to human musicality; musicality as a listener, our capacity to appreciate musical expression, and musicality as a player, our capacity to express music. From the perspective of musical interaction, the musicality of a musical interface could be considered as the degree to which the system is able to mediate musical expression.

1.8.2 INTUITIVE INTERACTION

The meaning of *intuitive* in the context of interaction has been the subject of some discussion, mainly because the term is used widely but until recently there has been no attempt at a formal definition. Raskin (1994) may have started the discourse on this topic. He proposed that *intuitive* in the context of interaction is almost an exact synonym of *familiar*, i.e. intuitive interaction makes use of our existing, transferable skills.

Bullinger, Ziegler, and Bauer (2002) explore intuitive interaction further, with the purpose of finding principles to design for intuitiveness. They suggest that intuitive interaction could be interpreted as *immediate usability*; this may only be achieved in the case of simple systems, but *immediate usability* also suggests a general goal of minimising the gap between a user's capabilities and those required to use a system, and leveraging existing skills.

Turner (2008) proposes two '*distinct but overlapping*' meanings of intuitiveness; intuitiveness as familiarity and intuitiveness as embodiment, the latter being more relevant for tangible systems that use a *grasp and manipulate* approach, where perceptual-motor skills are a key part of interaction. Exploring intuitiveness and embodiment further, Antle, Corness, and Droumeva (2009) looked at the role of embodied metaphor. They found that the use of embodied metaphors in the mappings of an interface could support intuitive interaction.

A definition of intuitive interaction is proposed by Naumann, Hurtienne, Israel, Mohs, Kindsmüller, Meyer, Hußlein, and Research Group (2007):

'A technical system is, in the context of a certain task, intuitively usable while the particular user is able to interact effectively, not-consciously using previous knowledge.'

Although this is a useful definition in many HCI contexts, there are reasons why this definition is not so useful for defining intuitive interaction with musical interfaces. Firstly, the definition is task oriented, whereas musical interaction, and creative interaction in general, is not necessarily so. Musical composers tend to exhibit a cycle of ideation and then evaluation (Coughlan and Johnson, 2006). While there is a wider goal of creating a piece of music, this process can be explorative, as opposed to a process of completing a set of specific tasks. The definition also talks about being able to *interact effectively*, and again this is not necessarily part of creative interaction. Interaction does not have to be effective for the creation of art, and a good example of this would be the practice of circuit bending (Ghazala, 2005; Collins, 2009), where electronic devices are deliberately dismantled, modified and repurposed in order to achieve sonic results unintended in the original design. A circuit-bent children's toy may not

be something that can be used effectively, although it may certainly be artistically or musically inspiring. Studies of computer musicians show that constraints and resistances in interactive systems can provide creative inspiration to their users (Bertelsen, Breinbjerg, and Pold, 2007; Magnusson and Mendieta, 2007), as can unpredictability (Gelineck and Serafin, 2009b). So, in fact, quite opposite to the definition, *ineffectiveness* can play an important role in musical interaction, at least in the context of composition.

It would seem that with the nature of musical interaction being so open ended, a less specific working definition is required. Returning to Bullinger's interpretation of intuitive interaction as *immediate usability*, in the context of this thesis, intuitive interaction can be considered to be *immediate musicality*.

CHAPTER 2

MUSICALITY AND DIGITAL MUSIC TOOLS

‘Tubby and his crew treated the studio
itself as a musical instrument.’

Greg Milner (Milner, 2009)

2.1 Introduction

The wider motivation for this thesis is to look at ways of enhancing the experience of interaction with digital music tools for computer musicians. This chapter looks at where typical problems can occur, the factors that cause them, and outlines typical systems that experience these problems. In doing this, an area of focus is carved out, establishing where problems need to be solved, along with the nature of these problems.

To begin this task, I will look at the musical activities of composition and performance. Conventionally, these two terms are considered to refer to separate activities, but when considered in the context of digital music, we shall see that there is much less difference between them, they are in fact part of the same process. I shall then examine the tools commonly used by computer musicians, and look at ways in which they are used. While we have instruments that are designed specifically with performance in mind, tools intended for composition and production work are rarely designed with the same considerations for interactivity, despite there being a strong performance element to their usage. I will put forward a holistic perspective of computer music tools, that considers tools aimed at composition and production to be instruments in their own right. Following on from this, we need to treat interaction design for these tools with similar considerations as for performance instruments; designing for embodied interaction is the key to this.

2.2 Interaction in Composition and Performance

Composition and performance are conventionally considered to describe two different activities, but in terms of computer music they are really integral parts of the same process. It's necessary to clarify this now, as later it will be important to consider what a composition tool is, and what a performance tool is.

When discussing composition, its meaning must be considered in the context of the instrument or system being composed for. As an example, consider first composition without computers, and the case of composing for an orchestra. The composer is physically detached from the instruments they are composing for, and detached from performance of the music. Although one may have a good concept of what the individual instruments sound like, it may be difficult to tell exactly how they would sound together when played. Composition in this case is internalised, and interaction happens mostly between the composer and the score. When composing for a solo instrument such as the piano, the process would be quite different. With the possibility of hearing the exact outcome of the composition, and also with the ability to try out or improvise sections of the piece as they are being written, composition becomes in this case a much more interactive experience.

Looking at electronic and computer music composition, again the nature of composition is closely tied in with the medium. One important element that sets computer music apart is the additional complexity of unexpected sonic outcomes. Early composers using punched cards to program computers would have had an even more detached experience than an orchestra composer. There was no real-time audio and possibly no idea of how the program

would sound, given that they were designing the instruments as well as the sequences. A musician using a piece of modern sequencing software however has a much more interactive experience; they can preview sounds and sequences as part of the editing process, usually in real-time, and listen to a piece as it evolves. Because of the potential for unexpected sonic outcomes, computer music composition is greatly enhanced by an interactive approach; sound design is a fundamental part of the composition process. Rather than being prescriptive, it's fundamentally an exploratory, experimental and interactive process.

It can be seen that in the context of specific mediums, the composition process can include elements of improvisation or performance. In order to understand these performative elements more fully in the context of electronic and computer music, I would like to explore some examples where composition and performance merge.

2.2.1 WHERE COMPOSITION MEETS PERFORMANCE

Moog, in a 1967 article describing the current state of electronic music technology (Moog, 1967), emphasised the performative elements of composition in the electronic music studio. Discussing new trends, he comments on the inefficiency of the process of tape manipulation in classical studio composition, contrasting this *programmed composition* with what he termed real-time performance in the control-voltage based studio. Chadabe (Chadabe, 1984) makes the composition-performance link more explicit, in what he termed *Interactive Composing*. He describes this process as creating an audio system that will react to input from a performance device, and then simultaneously composing and performing by interacting with the system. The audio system should not react in an entirely predictable way, creating a dialogue between itself and the musician. An in-depth account of the experience of interactive composition is given by Palacio-Quintin. She played a sensor-augmented flute (the *hyper-flute*) with an interactive system built in Max/MSP (Palacio-Quintin and Zadel, 2008). She comments on how users of such systems are simultaneously composers, performers and improvisers, and on how it's difficult to separate composer and performer with these systems. Turntablism is another area where these areas combine, where DJs use the combination of turntable and mixer as an instrument. Smith comments on the role of the DJ as composer and performer: '*The division between performer and composer is rarely an issue in DJ genres as the DJ combines both roles - composer in the selection of tracks and their restructuring via a number of compositional processes via turntables to create a new track, and performer in the performance of this process in the club environment.*' (Smith, 2000).

This scattering of examples demonstrate how the merging of performance and composition can be observed in a number of facets of electronic music. Studies of computer music composers also reveal more about this relationship.

2.2.2 COMPOSER STUDIES

Bertelsen et al.'s (2007) case study of two computer music artists makes some interesting findings on the nature of interaction in the composition process. They carried out a case study of two computer music composers in Denmark, with the aim of exploring the way that computers are used in creative work. They observed that the musicians' computers and associated

technology fulfilled a number of simultaneous roles for them, including that of composition tools and instruments for performance. During use, these roles changed rapidly, the musicians' focus shifting between the sonic output and the objects creating the sound. This focus shifts frequently so that the distinction between these roles is hard to clarify; the acts of composition and performance become indistinguishable.

Bertelsen et al. proposed that the musicians' software presented several instrument-like characteristics. The software is playable, it presents controls for manipulating sound in real-time. Further to this, the algorithms have unique sound profiles, making each one like a different instrument. The artists in the case study understood that they needed a disciplined approach to learning their software in order make the most from it creatively; like an acoustic instrument, they acknowledged that virtuosity was not easily achieved. They summarise this as follows:

'... the software is comparable to a musical instrument since the software becomes the object of his attention and something he explores, tweaks, observes, and challenges in a continuous shift of focus between the sounding output and the instrument.'

In a study of cognitive styles in the use of software for electroacoustic composition, Earglestone, Ford, Holdridge, and Carter (2008) collected questionnaire responses from 18 composers, revealing aspects of their work processes. Some of the composers describe how improvisation was part of their workflow, either in a fine-grained manner of improvising small elements, or with a larger scale approach of improvising longer sections and editing them down. Some revealed an emergent approach, the process taking place as an interactive dialogue between the tools and the musician, mediated through different tools or source materials.

These two studies are examples of how on a high level, composers may use performance as part of the composition process. Furthermore, in terms of low level interaction, the artefact used to create music doesn't have to fulfil a distinct role of tool or instrument, but can fulfil both roles.

2.2.3 COMPOSING INSTRUMENTS

The act of composition in the context of electronic music also merges with the act of instrument design and creation. As Magnusson (2009) says, the difference between composer and performer is non-existent, *'the act of composition dissolves into the act of instrument design, and often performance as well'*. To give an extreme example, take the practice of livecoding (Collins, McLean, Rohrerhuber, and Ward, 2003; Sorensen and Brown, 2007), where an instrument can be designed programatically in front of an audience as part of a musical performance. Livecoders often consider the programming language to be their instrument (Blackwell and Collins, 2005; Wang and Cook, 2004). This can also be seen in Chadabe's interactive composition, where the first stage of composition is to create a musical system. Both composers in Bertelsen et. al's study talked of how configuring or building the instrument was part of the creative process. They manipulated software systems in an exploratory way, shifting focus between instrument and sonic output as part of the wider process.

2.2.4 MUSICAL ACTIVITY

To summarise, the terms composition and performance conventionally refer to separate concepts, Bown, Eldridge, and McCormack (2009) call this the *acoustic paradigm*. However, by examining studies of users of electronic and digital music tools, and looking at real-world examples of musical practice, it can be seen that, in electronic music, composition and performance are in fact two facets of the greater whole, which is musical activity. The act of composition subsumes the acts of instrument design, improvisation and performance. Tools intended for composition or functional use, for example the turntable and mixer, can also be performed with as instruments; however, as shall be demonstrated, there are facets of interaction that may preclude this mode of use.

2.3 Modes Of Interaction

Anecdotally, composition is sometimes talked of as slow performance. In considering the nature of composition and performance, we have seen technology intended for functional or compositional use appropriated for performance. To analyse the relationship further, it's useful to focus on modes of activity; when can musical behaviour be considered performative and when can it be considered as functional or compositional?

Magnusson suggests that digital music systems can be divided into those intended for use in *productive* or *expressive* contexts (Magnusson, 2009). A production system is generally not designed for realtime use, while an expressive system is intended for performance and realtime interaction. While there are systems that can be placed comfortably in either of these groups, many systems also exist in a grey area between the two. Take for example a typical Digital Audio Workstation (DAW) application. Interaction may be functional, for example searching for samples on a hard-drive through a series of dialog boxes. It may also be performative, for example playing a softsynth with an external MIDI controller. It may also be somewhere in between, for example switching through plugin windows and adjusting parameters while a sequence is playing. The DAW contains elements which are intended as either productive or expressive; as we have seen with Bertelsen's study, these contexts blur in rapid use. The DAW can be seen as a gestalt collection of both types of elements such that when in use, musical activity will vary on a continuum between productive and expressive. This can also be observed beyond DAWs; activity in electronic music rarely consists of interaction with a single productive or expressive artefact, it typically takes place with a collection of artefacts from each context, or artefacts whose use entails both modes of activity.

Modes of interaction in digital music are one focus of Hunt and Kirk's study of mapping strategies (Hunt and Kirk, 2000). They contrast two modes of thinking that, they suggest, are involved in musical interaction: *analytic* and *holistic*. Analytic thinking describes when attention is focused on decoding information towards a specific goal, with a sequential thought process. A holistic mode of thinking describes a wider, sub-lingual awareness of the environment, where many things may be perceived at once. Holistic thinking focuses on the wider situation while analytic thinking focuses on detail. Hunt and Kirk propose that one of the problems of

interaction with digital music systems is the use of choice based interaction sequences, which force the user into analytic thinking mode when holistic thinking is more suitable for musical interaction. They propose the concept of *performance mode* interaction which is more suited to holistic thinking. In performance mode, the user can interact with a system in a continuous, parallel and realtime way, rather than in a non-realtime serial dialogue. This mode of interaction is closer to the way in which musicians interact with acoustic instruments. They also highlight the difference between performance and what they term *fast analytical editing*, which describes analytic mode interaction being carried out fast enough to constitute a performance. They consider this to be completely different from creative performance, noting that editing can be performative in a case where the interface supports explorative, continuous modes of interaction.

Drawing these strands together, Hunt and Kirk's holistic mode of thinking roughly aligns with Magnusson's expressive mode of interaction. We are then left with productive mode, which can be holistic and expressive given that the correct modes of interaction are available; it may also be analytic when the interface precludes holistic interaction. In the world of digital music, an instrument will typically be a gestalt collection of elements that afford varying modes of interaction. In use, the musician can switch between these varying modes, and given that the interface supports holistic interaction, the acts of *production* and *expression* can become indistinguishable from the greater whole. Interfaces that force the user into analytic mode can preclude holistic mode interaction, though redesigning these processes to support Hunt and Kirk's performance mode could solve this problem.

2.4 Computer Music Instrumentation

2.4.1 THE ANALOGUE STUDIO AS AN INSTRUMENT



Figure 2.1: King Tubby in his studio. ©Dennis Morris www.dennismorris.com

Having established that composition and performance are elements of the same process, I would like to focus on some explicit examples of electronic music tools designed for composition and production work being used as instruments. I will begin by stepping back and looking at aspects concerning interaction in the *studio-as-instrument* music production aesthetic. Having looked at this, I will continue by examining how this aesthetic fits into the current digital world.

The quote from the chapter heading continues as follows:

"Tubby and his crew treated the studio itself as a musical instrument. 'Everything was live,' Smart says. 'All the drops and everything was done live over and over again. I'd be cutting twenty, thirty, or even forty dubs in a day. It was all improv, because we had the vibe, the feel. We didn't even have to think about it. You would hear [the finished dub] and go 'Whoa, I didn't even know I did that one.'" (Milner, 2009)

This is a classic illustration of how some artists and producers in the popular music domain, beginning in the sixties when technology opened up new possibilities, began to perform not only with their acoustic instruments but also with studio equipment. They used it as both a musical instrument and a compositional tool (Collins, 1993). This blurred the traditional boundaries between composer and performer, the process of composition changing from a prescriptive process to one of exploration and interaction with tools (Eno, 2002). This blurring was not only in terms of musical role but in terms of the way people interacted with studio technology. King Tubby (see figure 2.1) was an innovator in this respect, developing techniques for using customised studio equipment to improvise music (Veal, 2007), and spawning *dub*, a new genre of music and the precursor to modern remix culture (Toop, 2002). He, along with his crew, became in effect virtuosic at playing their particular collection of equipment. They created music with their machines, interacting with them intuitively and expressively, perhaps as you would with an acoustic instrument.

The music press is replete with anecdotal references to the *studio as instrument* aesthetic, from the late sixties until the present day (e.g. Lee, 1997). Another compelling example of this style of interaction is DJing and turntablism (Smith, 2000), where again, artists use a collection of machinery as an instrument, giving virtuoso performances (Pinch and Bijsterveld, 2003) and interacting with a level of skill, subtlety and expression you might associate with acoustic playing. Ihde suggests that this use of machinery reflects trends in modern science where the new scientific instruments are, for example, colliders and particle collectors, *played* by scientists. In parallel, musical instruments can now be '*large, complex, high tech*' (Ihde, 2007).

2.4.2 THE DIGITAL MUSIC STUDIO AS AN INSTRUMENT

Stepping into the digital world, the 'studio-as-instrument' aesthetic continued; while some new equipment restricted the user interface and abstracted operation, other designs continued with the scientific instrumentation paradigm of the analogue interfaces, allowing more fluid interaction in comparison. The digital studio could be portable, allowing artists to set up a collection of equipment on stage and perform with it. A prominent exponent of this practice was the group Orbital (figure 2.2). The duo used a technique of improvising arrangements with Alesis MMT-8 looping MIDI sequencers, while altering sound parameters using hard-



Figure 2.2: Orbital Live, 2010 ©Scott Sanders ssandars@inet.net.au

ware synthesiser controls and a mixing desk (Buskin, 2006). The following interview extract illustrates their views about this performance technique (Détourn, 2001):

[But to play devil's advocate for a moment, what can really go wrong when you're just tweaking a few knobs, buttons and sliders?] "Fucking anything," Phil maintains, "and if it's a piece of piss, why isn't everybody else doing it? If it's so easy why aren't there loads of Orbitals out there?" "Technically speaking," Paul asserts, "jamming with sequencers isn't that difficult. Although I'd like to see someone do it with the proficiency that I've developed over the last ten years. I know my MMT-8s inside out. I know how to punch sequences in half-way through the bar and get breakbeats to do different things. The fun of it is in the improvising. But it's not how difficult it is - it's about how entertaining it is. We use the same sequences every night but it's a different arrangement. I'm not much of an instrument player. We jam with the sequencers. The thrill for me on stage is arranging everything live. I can throw the loops in and keep it going as long as I want, [blending] elements in many different ways." "We feed off the audience" Phil expands. "We can try to whip them up into a frenzy or we can go dub stylee. We've got the mixing desk, different effects, all the synths up there that are having MIDI signals sent to them, and the amount of manipulation of the sounds that you can do as it's going are amazing. It's a big organic quagmire." Paul delivers the final word on the matter: "You've got to be creative to be able to arrange the track, because if you're not, it's going nowhere."

Another example of this aesthetic is the live setup of electronic music act *Mouse On Mars* (Kirn, 2012). The trio describe themselves as an augmented band, playing with a combination of laptop software, midi controllers, hardware synthesisers, acoustic drums and vocals. Each

musician plays on a different set of devices and instruments, and the audio signal is passed around between the three players so they can all modulate each others output. Andi Toma comments on how they improvise with this collection of music equipment:

‘... so we have the basic structure and we try to be as free as possible to play with it, I mean it’s kind of a dangerous setup so always things can go wrong but this is the good thing’

Current DJ practice also incorporates digital studio technology in a similar way, blending the mixing of tracks with other elements generated and controlled by a range of studio tools and interfaces. Lopes, Ferreira, and Pereira (2012) describe the use of multitouch controllers such as the Lemur, and the use of hybrid setups which might include analogue turntables, analogue mixers, laptops, digital sound generators and digital controllers. DJs are using tools varying from standard MIDI controllers, to more esoteric interfaces such as the Wiimote (DJ WiiJ, 2010) or ReacTable (Hansen and Alonso, 2008).

Again, these are examples of audio hardware being used with musicality. The digital devices these musicians perform with on stage tend to follow the interface design of analogue equipment; they present individual controls directly to the player, encouraging fluid interactivity. Other digital tools however do not allow the same quality of interaction, presenting a greater degree of abstraction and lesser degree of control to the user. This contributes to what Armstrong (2006) terms *the disconnect*, between performer and audience, and performer and instrument, referring to an increasing sense of disembodiment between musician and music. While Armstrong focuses on this disembodiment in relation to performance practice, I would like to explore this further in relation to the use of tools intended for composition and production work.

2.4.3 DISCONNECTION

The very nature and possibilities of mappings in electronic music move the musician away from the direct musician-instrument relationship of acoustic playing. This was evident even before computer use was commonplace; Keane (1979), discussing the use of knobs and switches in the electronic music studio, describes a sense of disembodiment between composer and music, giving the example that violent or forceful sounds can be created with little physical effort. He asks: ‘*does the lack of a direct relationship between the composer’s own energies, or vicariously experienced energies, and the energy of the musical result exert an influence on the effectiveness of a work in the electronic medium?*’. Since this period, computers have become increasingly ubiquitous in electronic music, and interaction between musicians and sound tools has becoming increasingly abstracted. Ryan (1991) describes the problems composers were experiencing with computers in the early 1990s:

‘Certainly in computer music the problem is not lack of form, it is the immense mediating distance which confronts each composer when encountering the computer. Despite twenty years of programming for music, the territory gained seems quite small compared with the empire of musical aspiration. Many composers long to regain some sort of musical spontaneity.’

In general, recent technical innovations have tended to reduce opportunities for musical interaction (Winkler, 1998). This progressive disconnection is illustrated well by Øritsland and Burr's history of interaction styles (Øritsland and Buur, 2000; Djajadiningrat et al., 2007), discussed here in relation to musical technology. They classified modes of interaction historically, belonging to four epochs:

Machine Cowboy, 1933-1969 This epoch is characterised by simple, geometric forms, strong forces and direct feedback. Musical examples from this era would be early recording and tone generation equipment, for example as pictured in figure 2.3.



Figure 2.3: Siemens Studio for Electronic Music, ca. 1956

©Ben Franske, http://en.wikipedia.org/wiki/File:DM_Recording_Studio.jpg

Analogue Professional, 1970-1979 Components are miniaturised, direct feedback mostly disappears and activation forces are reduced. Instead of interaction requiring bodily force, a single hand can now be used. Visual feedback to the user is increased, with LEDs and meters. Figure 2.4 shows an analogue synthesiser from this era, produced between 1975 and 1979.



Figure 2.4: A Roland System 100 Analogue Synthesiser

©Hannes Grobe/AWI, http://commons.wikimedia.org/wiki/File:Roland-System100_hg.jpg

Digital Hacker, 1980-1994 Digital electronics take over from analogue, and miniaturisation continues. The interface is separated from the functional part of machines; as LED displays become more common, control becomes more abstract and less tactile. Controls are accessed through hierarchical menu systems instead of one-to-one controls. The iconic Akai S1000 sampler (figure 2.5) is a classic example of this trend, where a relatively complex instrument is operated from an interface consisting of an LED display, 32 buttons and two rotary dials. While performance parameters can be controlled from external MIDI controllers, operation of the machine itself takes places purely through this limited interface.



Figure 2.5: An Akai S1000 Sampler

Molly, 1995 onwards Øritsland and Burr describe how in the world of product design, this epoch signifies a move away from square box devices towards direct control and manipulation.

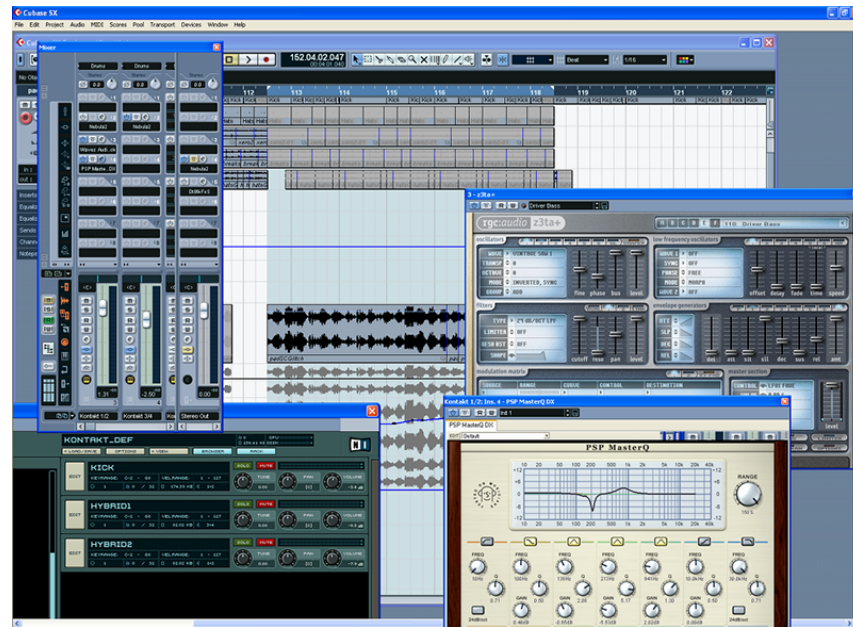


Figure 2.6: Steinberg Cubase SX, Digital Audio Workstation software

There are examples where the progression in music technology does not match in time with Øritsland and Burr's taxonomy, for example the recording studio pictured in figure 2.3 would contain elements that fit better into the *analogue professional* era. Music technology does however follow the general trend shown in the taxonomy, which demonstrates a progressive

move over time from physical and direct interfaces towards abstract and virtualised ones. The exception is in the last era: Molly. While industrial design moved away from abstraction, the design of electronic music interfaces stayed on the same course. As personal computers become increasingly capable at real-time digital signal processing, music technology moved away from the physical artefact and was virtualised (e.g. figure 2.6). The principal interface became the GUI, mouse and keyboard, along with accompanying performance controllers. The mouse, keyboard and GUI combination is particularly relevant in this chapter, from now on I shall refer to it as the *PC* interface.

Currently, hardware controllers to accompany the PC interface are widespread, and largely still based on the analogue era scientific instrumentation paradigm. Computer software meanwhile has become far more complex and multi-layered. This leads to an increasingly indirect mode of interaction between the user and tools; however, more recent developments in technology, DIY culture and human computer interaction are bringing physicality and directness of control back into computing.

2.4.4 RECONNECTION

2.4.4.1 RECENT DIRECTIONS IN ACADEMIC RESEARCH

The NIME¹ (New Interfaces for Musical Expression) (Poupyrev, Lyons, Fels, and Blaine, 2001) community is the group within academia most highly focused on issues surrounding musical interaction with computer systems. Researchers in this area are investigating new technologies and methodologies for creating expressive digital instruments, and trying to understand more fully the relationship between musicians and computers. Key issues in this field include gesture (Wanderley and Battier, 2000; Jensenius, 2007), pattern and sequence recognition (Cont, Coduys, and Henry, 2004), the use of novel sensors and actuators for creating new forms of musical interaction (Freed, 2008), the augmentation of acoustic instruments (Newton and Marshall, 2011), systematic approaches to instrument design (Ryan, 1991), the realisation of expressiveness (Arfib, Couturier, and Kessous, 2005), collaborative music making (Bryan-Kinns and Hamilton, 2009) and embodiment (Magnusson, 2009). Underpinning all of this is the subject of mapping (Pressing, 1990; Hunt, Wanderley, and Paradis, 2002), focusing on the myriad ways in which we can link the physical world to abstract digital structures and sound generation processes in order to create music. This issue is explored further in section 3.4.

NIME has grown in parallel with emerging fields such as tangible interaction, and physical computing (Igoe and O'Sullivan, 2004). Researchers are exploring the use of more direct, physical interfaces for computers, moving beyond the PC interface, and beyond the scientific tool paradigm. Ishii and Ullmer (1997) clearly set out the motivations and issues in tangible computing, proposing a '*seamless interface between people, bits and atoms*'.

A prominent example embodying the principles of NIME is the ReacTable (Jordà, Geiger, Alonso, and Kaltenbrunner, 2007; Jordà, 2009). The ReacTable allows the collaborative control of a modular synthesis system in a tabletop environment, with users interacting by moving and twisting objects on the surface and also by using fingertip control. The ReacTable acts

¹<http://nime.org>

as a controller for underlying synthesis software such as PureData or Max/MSP. This software would typically be operated with a PC interface. Introducing direct physical control into the manipulation of patches with ReacTable promotes complex, skilled, expressive and explorative use, bringing a strong performance element into interaction with the software.

2.4.4.2 MULTI-TOUCH AND MOBILE TECHNOLOGY

Multi-touch provides a more physical and potentially expressive and intuitive way of working with computers than the conventional PC interface approach. The combination of accelerometers, cameras and other sensors present in many mobile multi-touch devices creates a recipe for some compelling musical interfaces, and ample opportunities for software developers to exploit bodily interaction. Affordable multi-touch computers such as the iPhone and iPad have already had a significant impact in the world of computer music. JazzMutant's *Lemur* music controller has already shown that this style of interaction is popular, and being actively used by professional musicians. Artist Richare Devine discusses his experience of using a multi-touch system for music production (JazzMutant, 2010):

‘You can’t control software like you can with the Lemur; the closest is with Max/MSP but not with the same elegance and tactile interaction that is so important. I can feel the manipulation and I’m shaping the sound almost like a sculpture piece because you have such a tight connection and level of control, it’s brilliant!’

This controller has since been eclipsed by the more affordable iPad series, which has seen a flood of applications created for it from both small developers and established music software houses. With other tablet systems planned for release in the near future, multi-touch’s more direct mode of interaction is having a significant impact on the way we make computer music.

2.4.4.3 DIY MUSIC CULTURE

With a recent surge in DIY computer music culture, fuelled by low-cost interface boards such as the Arduino (Banzi, 2009), we are seeing some compelling new interfaces being created by hobbyists. Revolving around blogs such as *Create Digital Music*² and globe-trotting events like *Music Hackday*³ and *Handmade Music*⁴, amateur hardware hackers are using cheap sensors and repurposed hardware to create some extremely inventive interfaces. Projects such as these expand the pool of interfaces and instruments the computer musician may draw on as part of the computer music studio, and the use of novel sensors and interaction techniques expands the performative potential beyond conventional computing paradigms.

2.5 The Body In Interaction

Having examined the trajectory of interaction style in electronic and computer music, physicality has been a recurring theme. A clear understanding of the relationship between mind,

²<http://createdigitalmusic.com>

³<http://musichackday.org/>

⁴<http://handmademusic.noisepages.com/>

body and environment is at the core of understanding the nature of interaction between musicians and music technology. Interconnected research in fields such as cognitive science, philosophy of mind, artificial intelligence, robotics and situated cognition is now giving us a deeper understanding of this relationship. At the core of this research is a rejection of the Cartesian mind-body separation, and instead the adoption of a view of humans as embodied agents in the world. Mind, body and environment can no longer be considered independently; *'knowing is inseparable from doing'* (Brown, Collins, and Duguid, 1989).

The \$50 million CYC system is an often quoted example that demonstrates this inseparability. CYC is an artificial intelligence project, deeply rooted in the GOFAI (good old fashioned artificial intelligence) tradition, that attempts to symbolically encode human knowledge. One of the aims of the project was to create an artificial mind. While CYC is an excellent example of an expert system, this goal has not been reached. Clark (1997) posits that the system is fatally flawed, due to the lack of any adaptive response between CYC and the real-world environment; there is no coupling so it cannot make sense of real-world problems. Without the grounding of its representations, the knowledge encoded in the system relies on human interpretation to give it meaning (Anderson, 2003).

Clark and Chalmers (1998) proposed the idea of *active externalism*, which describes how we use the world as part of our cognitive processes, with mind and environment acting as a coupled system. In his book *Being There*, Clark (1997) describes the *leaky mind*, which escapes the confines of the body, using the environment for support. An example of this would be using a book mark to remember your place in a book rather than remember the page number, or using our perceptual-motor skill and the physical properties of an abacus to perform mathematical calculations rather than perform them internally. At the moment, the words and concepts that will form this sentence are spread across several nodes of an on-screen mind-map, created by mind-mapping software which I am using to spatially organise my thoughts; arguably this software is an extension of my cognitive process. Kirsh and Maglio (1994) proposed that we should distinguish between pragmatic and epistemic action. Pragmatic action happens when we deliberately change the world, for example grating a piece of cheese. Epistemic action describes when we use the world to enhance cognitive function, for example using the abacus. The concept of epistemic action and the ideas surrounding the philosophy of embodiment have provided fuel for a new way of thinking in the field of human-computer interaction.

2.5.1 EMBODIMENT AND HUMAN-COMPUTER INTERACTION

Dourish (2001) combines thinking on embodiment with HCI in his book *Where The Action Is*. He suggests that embodiment should be the central concern in interaction design, with current computer interfaces failing to take account of the ways humans are coupled with the world. Dourish details a history of interaction and, like Øritsland and Burr, divides this history into separate epochs:

Electrical This was the era of analogue and very early digital computers. The boundary between software and hardware was not easily delineated, and reconfiguring or programming a device required a detailed knowledge of its design.

Symbolic Programming a computer now requires less knowledge of specific hardware, and instead knowledge of more generic concepts such as registers or accumulators. Programming moves from numeric form towards higher-level forms such as assembler language or LISP, although code is still input on punched cards or similar media. Programming is the main way of interacting with a computer, and this more symbolic level of interaction is more natural and intuitive than electrical era interaction.

Textual This marks the first form of the *interactive loop* between human and computer; a dialogue forms as the computer accepts textual commands from the operator and returns textual feedback.

Graphical Many new possibilities for interaction are opened up with the move from one-dimensional text streams to two-dimensional graphics. The task of managing information is now spatial, this enables the exploitation of further aspects of natural human skills to enhance the experience of interacting with computers, including peripheral attention, pattern recognition, spatial reasoning and visual metaphors.

Through these epochs, a trajectory is clear where human-computer interaction progressively exploits more and more of our innate skills and abilities.

Dourish outlines what he sees as the next epochs of interactive computing:

Tangible and Social Approaches to Computing This reflects current trends in interaction research. He identifies three trends in tangible computing, firstly the distribution of computation across physically separate devices which may be aware of their location and other devices, for example printers. Secondly, the augmentation of everyday devices with computing power, such as toys, and lastly, the use of real-world objects to interface with computers, providing a more direct mode of interaction. Social computing focuses on applying an understanding of how humans interact with each other to improve interaction between humans and computers (for example Gill (2007) examines entrainment and tacit communication between humans, and lays the basis for a design framework for affording these qualities in human-machine interaction).

Embodied Interaction This is a new perspective on interaction that draws on commonalities between tangible and social computing, emphasising the importance of the way we interact with the everyday world. Dourish describes this as *'an approach to the design and analysis of interaction that takes embodiment to be central to, even constitutive of, the whole phenomenon'*. He sees this paradigm as a potential solution to many of the current problems in HCI.

The influence of thinking on embodiment has become more prevalent in HCI, now with a growing research community dedicated to tangible computing centred around the TEI (Tangible, Embedded and Embodied Interaction) conference⁵. In the closely related world of product design, this thinking is having similar influence. Djajadiningrat et al. (2007) explored the role of movement in human-product interaction design, emphasising how the body is currently neglected, and advocating an embodied approach taking advantage of human perceptual-motor skills. They make a strong argument for a move away from interfaces with narrow bandwidth binary controls towards multiple concurrent continuous controls. Increasing the degrees of freedom of control in this manner allows more complex interaction that can exploit our natural capacity for dexterity and expressiveness of motion. This multiparametric approach will be discussed in more detail in section 3.4.2.

⁵<http://www.tei-conf.org/>

2.5.2 EMBODIMENT AND COMPUTER MUSIC

Embodiment philosophy is also having an influence in the world of computer music. Armstrong (2006) outlines an *enactive approach to digital instrument design*, centred around embodied modes of interaction. He highlights the importance of embodied coupling between human and instrument, and lists five criteria of embodied activity:

1. Embodied activity is situated. It happens through the context of an agent's interactions with the environment.
2. Embodied activity is timely. The agent must act within realtime constraints.
3. Embodied activity is multimodal. The activity occupies a large part of the agent's sensorimotor capabilities.
4. Embodied activity is engaging. The agent is intimately involved in the task, giving rise to a sense of embodiment.
5. The sense of embodiment is an emergent phenomenon. It arises incrementally through a history of interaction between agent and activity.

Armstrong introduces the concept of the *computer-as-it-comes*, describing a generic personal computer, based on conventional HCI guidelines. The computer-as-it-comes precludes the conditions set out for embodied interaction, and this greatly reduces its potential to realise musical performance. Underpinning his approach is an emphasis on the emergent properties of interaction; he differentiates between two classes of interfaces: functional and realisational. A functional interface is task oriented, and has a finite set of possible interactions. The realisational interface is non-deterministic, allowing for the continual development of new encounters, echoing Paine's model of conversational interaction (Paine, 2002). A realisational interface, in opposition to conventional HCI goals of interface transparency, is also *present*, offering resistance to the user and thereby prompting engagement. These principles guided the implementation of a new musical instrument entitled *Mr Feely*. The hardware design of Mr Feely follows the scientific instrumentation paradigm, presenting a panel of knobs, buttons and joysticks to the player. Armstrong details two patches for the instrument which demonstrate enactive performance; one demands physical effort from the player to bring a rich musical response, and one invites the player to *surf* a fractal wave. The design of these patches and this system is rooted in Chadabe's concept of interactive composition, and it would seem that the emergent behaviour of the system has its root in the software design, although in performance one must consider the system holistically. The question arises as to whether these emergent properties could also be made implicit in the physical design of the controller, and what affect would this have on the quality of musical interaction? Is it possible to experience a more fine-grained form of emergence, present on a sensorimotor level? These questions will be relevant later on in this thesis.

While Armstrong believes that the computer-as-it-comes precludes the conditions of embodied interaction, Magnusson (2009) disagrees with this, stating that '*the computer clearly allows for an embodied interaction just like the acoustic instrument*'. He argues that in the case of screen based instruments, interaction can in fact be engaging, timely and situated, and depending of the software design, also emergent. The final condition, of multimodal activity,

is fulfilled by new interfaces which have been developed into ‘*sophisticated musical instruments*’. He gives the example of Nintendo’s Wiimote, which has been appropriated by musicians as a musical controller.

Given that both the *computer-as-it-comes* and *Mr Feely* might be classified as embodied instruments, in the analysis of embodiment in musical interaction, perhaps important factors are not only whether the conditions of embodied activity are filled but also the quality or degree of embodiment? Can we say that *Mr Feely* is more or less embodied than the *computer-as-it-comes*, and what does this actually mean in terms of musical interaction? Magnusson and Medieta’s survey of computer musicians provides evidence that embodiment is an important factor in the use of digital instruments. Their survey focused on the relation between acoustic and digital musical instruments and the body (Magnusson and Mendieta, 2007). Analysis of the 209 responses showed that musicians ‘*lament the lack of embodiment*’ in digital instruments, although it is also noted that it is impossible to create acoustic-style embodied interaction within the digital realm because of a lack of tactile feedback characteristics.

Returning to Hunt and Kirk’s modes of musical interaction (section 2.3) with an embodiment perspective, their *analytic* mode interaction could be considered as lacking the condition of timely interaction. Being forced into sequential challenge response mode lacks the condition of being timely, and precludes an embodied relationship with the interface, limiting musical interaction. Their proposed *performance mode* solves this by making the interface continuous and parallel, allowing timely interaction. The ReacTable is an excellent example of this style of interaction, where the commonly sequential elements of a process have been made into continuous elements of interaction. Jordà comments on this (Jordà, 2009):

In RTI [realtime interaction], time passes independently of the users’ actions, and the perception of ‘instantaneity’ is not as essential as the perception of ‘continuity’, primarily in time (but often also in space, as we will later explain). In this time continuum, the user’s actions do not need to wait for each system answer; like in a real conversation, there are no fixed turns and everyone is free to say anything at any time.

ReacTable is a clear example of a more embodied approach to interaction design enhancing the musicality of a system.

It’s clear that current thinking on embodiment is causing a paradigm shift in HCI and also in turn in musician-machine interaction. Designing systems with an focus on the exploitation of human perceptual-motor skill can have a pronounced benefit on user experience.

2.6 A Holistic Perspective of the Computer Music Instrument

Drawing these strands together, I wish to put forward a holistic view of the computer music instrument that de-emphasises intended use by the designer (e.g. production or performance) and focuses on the interactional affordances of a system. The instrument is considered to be

a gestalt collection of elements with varying affordances, perceived as a conceptual whole by the musician.

We first looked at the tasks musicians use their tools for, and saw that composition, improvisation and performance are far from separate concepts when considered in terms of computer music. Instead they can be seen as overlapping facets of a greater whole, which is musical activity. A holistic perspective of the computer music instrument views all tools equally in terms of musical activity and musicality; an artefact intended for production use such as a sequencer should ideally afford the same musicality in interaction as an artefact oriented towards performance.

We have also seen that the modes of interaction afforded by computer music systems can vary on a spectrum between productive and expressive. Given the right conditions of interaction, a set of tools can be used across these modes with musicality. However there are properties of interfaces that may preclude musicality, such as a lack of timely interaction, or a failure to exploit our perceptual-motor skills; in short, a lack of embodied interaction.

I would now like to map out the classes of computer music system that are the focus of this thesis, systems in common use where musical interaction is typically impacted negatively by interface design and lack of embodiment. I will do this by examining the quality of embodied interaction along with the conceptual complexity of the systems.

2.7 Problematic Areas In Musical Interaction

Figure 2.7 shows a variety of instruments mapped out according to two key factors in computer music interaction: embodiment and conceptual complexity.

2.7.1 DEGREE OF EMBODIMENT

We have already seen that the use of the body in interaction is key factor in the musicality of a system. In order to establish an approximate measure of body use, we can focus on key factors that affect this: the complexity of mapping and the degrees of freedom of motion afforded by the system. For the complexity of mapping, at one end of the scale we have one-to-one binary mappings such as buttons. At a higher level of complexity we might have linear scale mappings such as sliders or key velocities. At the highest level we could place non-linear, mixed divergent and convergent mappings, such as in acoustic instruments. Looking at freedom of motion, at the low end of the scale we can again put binary input such as buttons. At the highest level, we can again place acoustic style interaction, where there can be many degrees of freedom employed in the production of sound, and where subtle changes can make both subtle and large differences. These two factors are fairly congruous, so I would like to consider them under the same umbrella, which is the degree of embodied interaction afforded by a system. To be more specific, this can be defined as the degree to which the exploitation of human perceptual-motor skill is fundamental to the working of a system.

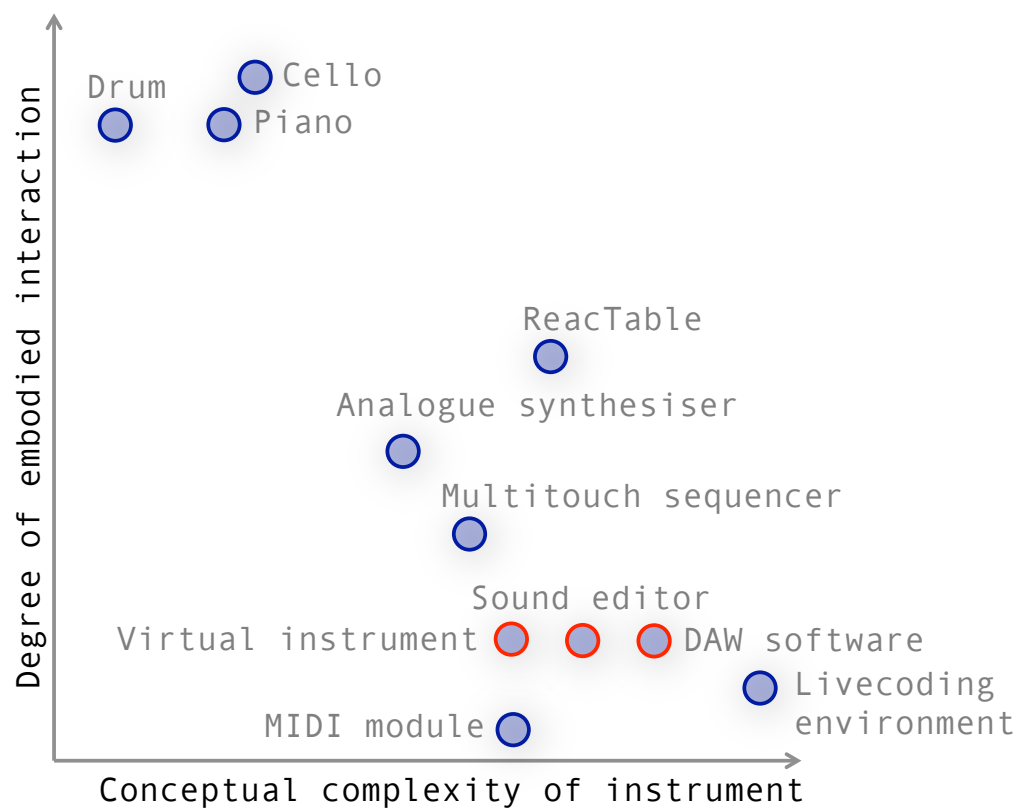


Figure 2.7: Embodiment vs Conceptual Complexity

2.7.2 CONCEPTUAL COMPLEXITY

Conceptual complexity is another key factor in interacting with a computer music system, and as with the level of embodiment, can be decomposed into several contributing themes. As has been demonstrated earlier in the history of product design, conceptual complexity is a modern issue; as technology has advanced, increasing amounts of functionality have been condensed into smaller and smaller interfaces. With this comes an increased distance between user and functionality, as increasing numbers of features are hidden behind decreasing numbers of controls. Important components of conceptual complexity are the level of abstraction in an interface, and the representational complexity of the system. A system at the low end of this scale may embody simpler functionality, such as playing percussive hits or playing notes. A system at the high end may embody a mix of high level concepts such as digital music theory, digital signal processing or computer programming.

2.7.3 ANALYSING THE CHART

This chart does not aim to be inclusive or precise. Rather, the intention is to map out some typical computer music systems in order to establish those where musical interaction can be enhanced, and therefore where this thesis focuses. Although gestalt instruments such as studios have been discussed already, these will be ignored for now for the sake of simplicity.

Acoustic instruments lie in the top left area of the chart, the percussion instruments first

with the piano representing higher complexity than the drum. The cello sits above the piano in terms of embodiment due to the greater affordances of a bowed instrument and the greater vibrational coupling to the body. Around the central area of the graph sits the analogue synthesiser, representing scientific instrumentation paradigm interfaces in general. Here, the controls are mostly continuous and linear, or discrete buttons. These elements afford lower degrees of freedom than acoustic instruments.

The MIDI module sits at the bottom end of all the systems in terms of embodiment. An example would be the Akai S1000 sampler mentioned earlier (see figure 2.5), where a limited set of mostly binary controls govern the complex processes necessary to run a sampler. A more extreme example can be seen in Korg's Wavestation SR, where a complex digital synthesiser engine with a large amount of parameters presents its primary interface as fourteen buttons and a two line LCD display.

At the far end of the complexity scale lie livecoding environments such as SuperCollider (McCartney, 2002), Impromptu (Sorensen, 2005) and Chuck (Wang and Cook, 2003). These systems represent the high end of conceptual complexity with their reliance on knowledge of computer programming and low level digital music theory. Their PC based text editor interfaces represent a more embodied form of interaction than the MIDI module.

The more GUI reliant systems sit above livecoding systems in terms of embodied interaction with their more prominent use of continuous input from the mouse, and also spatial and graphical elements.

More recent technological developments are represented by the multitouch sequencer and the ReacTable. Both of these interfaces improve the musical interaction limitations of PC interfaces. A multitouch system such as the iPad or iPhone affords a larger degree of freedom of motion than a mouse, and additional sensors such as accelerometers in handheld devices provide additional dimensions of control. The ReacTable creates a tangible interface for processes that would be operated with a PC interface, again allowing for a far greater freedom of motion.

Both the iPad and ReacTable interfaces provide improved versions of systems originally designed for personal computer use, and it's this same area that I focus on with the projects in this thesis. These systems are typically fairly high on the scale of conceptual complexity, but are limited in terms of musicality in interaction. The mouse and keyboard are a bottleneck, offering limited freedom of motion and only simple discrete and continuous linear mappings. The projects presented here are new input devices designed to interface with GUI software, but offering a higher quality of embodied interaction.

2.8 Conclusion

At the onset of this chapter, I stated the aim of building a picture of computer music interaction in order to establish in a general sense where typical problems of interactivity lie, and to understand what the causes are. A holistic view of computer music interaction has been presented, where instruments can be seen as collections of tools, affording varying modes of interaction and used for a variety of musical goals. This is a view supported by case studies

of composers, and by real-world examples of electronic music practice. The importance of embodied modes of interaction has been identified in two key ways. First, the property of timely interaction is key to a successful musical interaction. Secondly and most important, the more that an interface exploits human perceptual-motor skill, the better it can act as a musical interface.

I have presented a range of typical electronic and computer music systems, and established PC interfaces as an area where interactivity problems need to be solved, due to the lack of embodied control available for the control of musical tools on these systems. Given this, the aim of the projects in this thesis will be to look at ways of increasing the level of embodied interaction with new personal computer interfaces for the control of digital music tools. This will involve the development of new hardware, as well as the development of mapping techniques to interface these new interfaces with typical computer music software systems.

CHAPTER 3

DESIGN AIMS

‘Being a programmer I find it more interesting to find how these machines can do things they weren’t meant to do.’

Trent Reznor (Berger and Lengvenis, 1994)

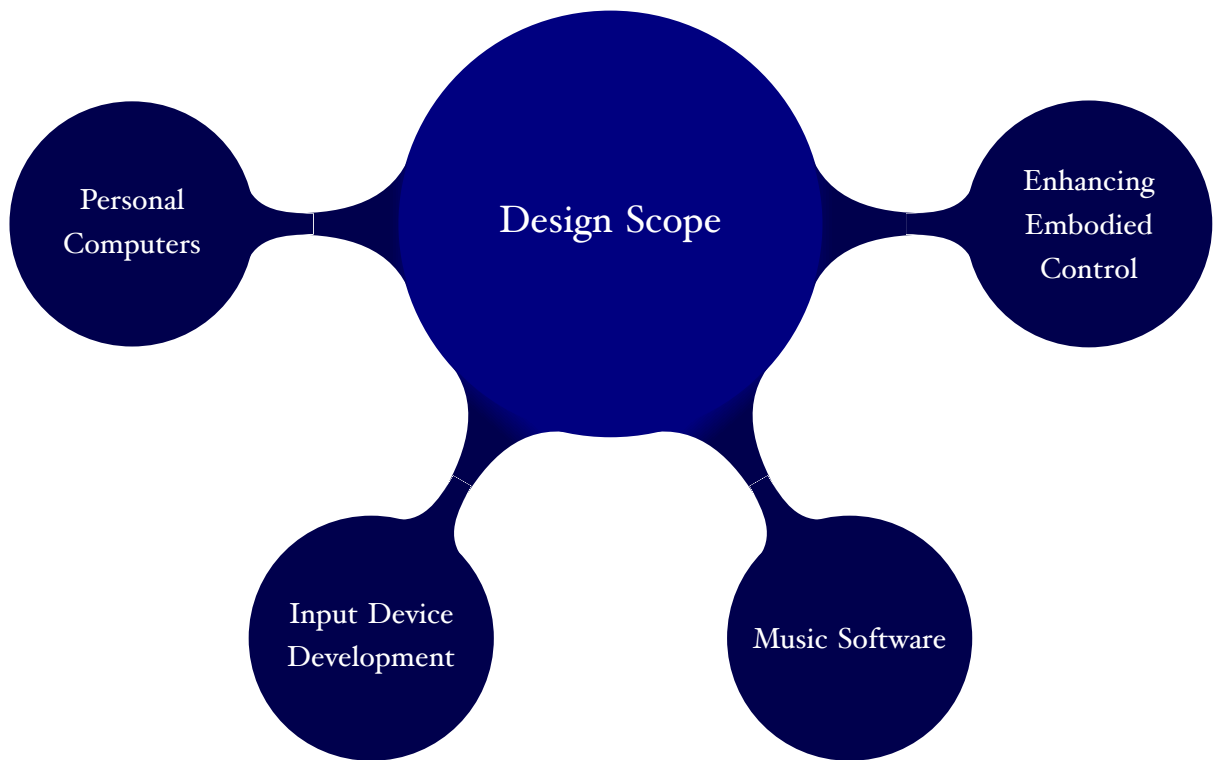


Figure 3.1: Design Scope

3.1 Introduction

Having set out the broad focus of this research (see figure 3.1), this chapter aims to establish global design goals for the creation of new musical interfaces within this area. By reviewing current academic research in the areas of creativity, digital instrument design and mapping techniques, design goals can be inferred to aid with the creation of new interfaces. Further to this, the philosophy behind the design process shall be explored, to help shed light on why certain design decisions were made in the projects described later. To begin with, I would like to zoom out and examine literature in the wider area of creativity research, to see how this can be applied to digital instrument design.

3.2 Designing for Creativity

3.2.1 DESIGN PRINCIPLES FOR CREATIVITY SUPPORT TOOLS

Shneiderman, Fischer, Czerwinski, Resnick, Myers, Candy, Edmonds, Eisenberg, Giaccardi, and Hewett (2006) report from a workshop which drew together a group of leading academic experts in the field of creativity support tools. One of the key outcomes from the workshop was a set of twelve principles for designing tools for creativity. These guidelines cover the wider field, including for example tools for writing and visual art, but are equally relevant for computer music and the creation of digital music interfaces. The design principles are as follows:

1. Support exploration.
2. Low threshold, high ceiling, and wide walls.
3. Support many paths and many styles.
4. Support collaboration.
5. Support open interchange.
6. Make it as simple as possible - and maybe even simpler.
7. Choose black boxes carefully.
8. Invent things that you would want to use yourself.
9. Balance user suggestions with observation and participatory processes.
10. Iterate, iterate - then iterate again.
11. Design for designers.
12. Evaluate your tools.

Reflecting on their suggestions, number two in particular is emphasised in Shneiderman et al's report. A low threshold will support novices, while a high ceiling allows mastery and sophisticated use. Wide walls support a range of possible usage scenarios. As will be seen, this guideline crops up repeatedly in various forms in other literature, having particular relevance to virtuosic play in music. Overall they suggest an open approach to designing for creativity, where an interface should inspire rather than prescribe usage. This is an approach I have attempted to follow in the projects described in this thesis.

3.2.2 FLOW

The concept of flow, put forward by the psychologist Csíkszentmihályi, describes an optimal state of heightened awareness when performing an activity; a state of effortless engagement, absorption and focus (Csíkszentmihályi, 1991). He describes flow experience as consisting of seven elements:

1. Clarity of goals and feedback
2. A high degree of concentration on a limited field
3. Balance between agility and challenge
4. A sensation of heightened control
5. Effortlessness of action
6. An altered perception of time
7. The melting together of actions and consciousness

The first three elements are pre-conditions of flow, and the last four describe the subjective experience of the flow state. Chen et. al clustered these elements into three categories: antecedents, experiences and effects (Chen, Wigand, and Nilan, 1999). The antecedent stage describes the preconditions for reaching the flow state. The experience stage comprises the merging of awareness and action, and the effects stage describes the inner experience, as the activity becomes autotelic.

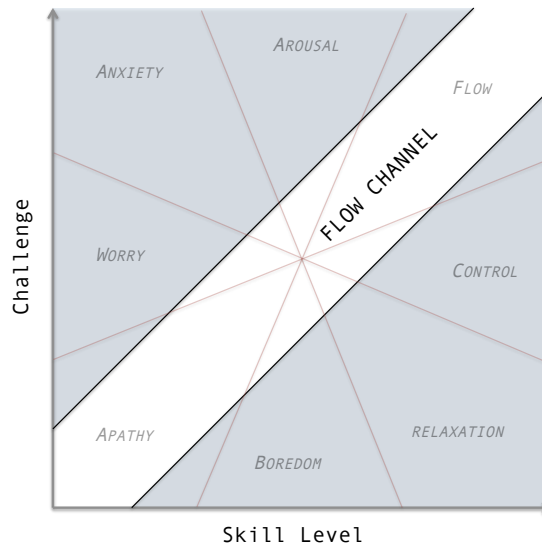


Figure 3.2: Flow: Challenge vs Skill

In terms of interaction, figure 3.2 illustrates flow as an optimal balance between the challenge of an activity and the skill required to carry out the activity. As skill level is increased, flow is lost if the challenge does not increase to match the skill level. Conversely, if the challenge is increased without a matching skill level increase then anxiety results and flow is lost. This relationship between skill and challenge, or between anxiety and boredom, may be explicit or implicit. For example, it is made explicit in some computer games where the challenge increases in discrete steps with each game level, to try and match the players' growing skill and therefore preserve entertainment value. For musicians, this relationship can be more implicit; in the learning of an instrument, a player will typically attempt to play more difficult material as they become more accomplished, and may become bored with playing music below their skill level.

Flow has a strong connection with creativity and music. This is demonstrated in MacDonald et. al's empirical study of creativity in group music composition (MacDonald, Byrne, and Carlton, 2006). They found that higher levels of flow in a group of students during a composition task correlated to higher levels of creativity and higher quality compositions.

Flow is also linked with musical performance; as Armstrong says (Armstrong, 2006):

'In short, the experience of flow, of a heightened sense of embodiment, involves an immediately palpable feeling of active presence in a world that is directly lived and experienced. Traditional though it may seem, these are qualities that I believe are central, and will remain central, to musical performance.'

His belief that flow is central to musical performance is backed up in research. Joann Marie Kirchner (2008) found that creating conditions that encourage flow experiences can help to alleviate musical performance anxieties. Musical performance has been used as a mechanism to induce flow; de Manzano, Theorell, Harmat, and Ullén (2010) used piano playing to induce flow experiences in their study of the psychophysiology of flow. A compelling account of flow as part of musical study is given by Burzik (2003), who proposes that practising in a flow state can have a positive effect on learning a musical instrument. He suggests a more playful, less goal oriented approach to practice sessions, creating conditions that encourage flow experiences.

Given the connection between musical performance and flow, it's important that any new interface for digital music should at least be able to create the conditions for flow states, and even better, should in some way encourage these states. There are two strong themes in flow, from which some design goals can be extracted. There is the important relationship between challenge and skill level. This is also related to Shneiderman et. al's second design principle, of low entry with a high ceiling; an instrument should be playable by a novice, providing little challenge when they have little skill. It should also be playable at virtuoso level, providing high challenge for a highly skilled player attempting difficult material. The other theme is that of timely interaction, which is a condition of embodied activity (Armstrong, 2006); non-realtime elements of an interface will interrupt a flow state.

3.3 Desirable Qualities in Musical Instruments

It has already been put forward that a high degree of embodied interaction is a very desirable quality in a *good* musical instrument. That is to say that a good instrument should make the most of human perceptual-motor skill in interaction. Further to this, what are other possible qualities that make a good instrument? Several studies from the NIME field make suggestions in this area (e.g. Cook, 2001; Wessel and Wright, 2001), from which more design goals can be drawn. A key text to refer to is Jordà's PhD thesis on digital lutherie, in which he dedicates a chapter to exactly this topic (Jordà, 2005, chapter 7).

Jordà proposes a set of design guidelines, based on academic research along with his own intuitive experience of instrument design. He attempts to establish a framework for designing good musical instruments. The key points from this framework are summarised below:

Challenge and Learning Curve

A good musical instrument should achieve balance between frustration and boredom; this directly reflects the need for facilitating flow states when playing an instrument. A good instrument will also have a learning curve that allows novice users to play the instrument, while also providing enough challenge and reward for virtuoso players.

Output Diversity

Diversity reflects the range of sonic outputs a musical instrument can produce. This can be viewed at three levels: micro, mid-level and macro. Micro-diversity describes subtle, nuanced changes in sound. Mid-diversity reflects performance contrasts, or how different two different pieces played with the same instrument can sound. Macro-diversity describes the stylistic flexibility of an instrument, the flexibility and adaptability of an

instrument to play varying styles of music in different contexts. It represents the instrument's potential to work within different musical roles and its compatibility with other instruments. The factor most important for achieving virtuosity with an instrument is micro-diversity.

Efficiency

Jordà defines this as follows:

$$\text{MusicInstrumentEfficiency} = \frac{\text{MusicalOutputComplexity} \times \text{DiversityControl}}{\text{ControlInputComplexity}}$$

This formula describes how an efficient instrument will have a good ratio of output complexity to control complexity; it will require less knowledge and effort to master. The diversity control term is included so that simple instruments with complex output will have a low score, for example a DAT machine where the operator just needs to press the start button.

Deep and narrow vs wide control

An acoustic instrument can be described as deep and narrow, that is to say that it enabled the musicians to control a small amount of parameters in a very nuanced way. Computer instruments have the potential of being too wide and too shallow; giving the musician limited control of a large amount of parameters. A good instrument should be well balanced in this respect.

Non-linearity

Most acoustic instruments exhibit some sort of non-linear behaviour, where the sound produced is not directly proportional to the state of the instrument at a particular instant in time, but instead depends on a history of the players actions. Jordà states that *'If non-linearity is at first intuitively seen as a source of potential uncontrol, it can therefore also mean higher-order and more powerful control'*. Non-linearity exists in a balance with the confidence a musician has in an instrument; while non-linear behaviour is essential, it should not inhibit a musician's trust in an instrument. Non-linearity can lead to virtuosity.

Predictability

A degree of randomness means that no two performances will sound the same. As with non-linearity, randomness can inhibit the development of skill at detailed control, and needs to be balanced with determinism.

Fault Tolerance

Many acoustic instruments are tolerant towards unexpected gestures and misuse, while digital instruments are generally fault-intolerant. An instrument that encourages misuse can inspire interesting experimentation.

Confidence

A musician needs to be able to trust an instrument in order to get the most from playing it, to explore the zones of non-linearity. Confidence in an instrument means a musician can truly make music with it.

Explorability

New digital instruments are self-taught; there is no standard repertoire, and lessons are not generally available. Therefore, a degree of explorability is essential for the new-comer to be able to learn an instrument, through improvisation and experimentation. While traditional instruments are generally deep and narrow, favouring a sequential learning process, new instruments are generally wider and lend themselves to parallel exploration.

Timing

Control granularity describes the rate at which musical parameters can be modified. An acceptable granularity is every 10-20 milliseconds for the most time sensitive parameters. Latency describes the time between a control action and a sonic result, research shows that 20-30 milliseconds is generally tolerable.

Sharpness and Inertia

Musical inertia is a high-level measure of musical timing, describing the speed at which an instrument can convey natural musical changes and contrasts. Jordà suggests that a good value for this is between 250 and 1000 milliseconds.

Expressiveness

A good instrument should be expressive, in that it should be capable of letting the player convey their feelings and ideas musically. Fine-grained control is a key ingredient for expressiveness.

To expand further on Jordà's framework, Gelineck and Serafin (2009b) shed further light on what makes a good instrument. They conducted an empirical study of the composition process of 18 electronic musicians. Their results reveal show the artists as enjoying an explorative approach to creating music; they enjoyed seeing where the process of discovery took them. The artists preferred a degree of uncontrol and unpredictability as part of the creative process. Gelineck and Serafin highlight explorability as an important quality of an instrument used for composition, and they suggest some design proposals to try and encourage this:

1. Design for unintended use.
2. Design for balance between an intuitive tool and an unpredictable tool.
3. Restrict the possibilities of the musical tool.
4. Make the tool compatible with everything else.
5. Give the tool a possibility of passing sound through it.

These proposals add further weight to several studies already discussed in this thesis. Designing for unintended use (see section 4.1), although seemingly paradoxical, could improve the quality of an instrument. This translates to designing instruments and interfaces that are open in nature; this crosses over with Jordà's concept of tolerance, being ready for misappropriation. Creating a balance of intuitive control and unpredictability is also reflected in Jordà's framework, in the balance of non-linearity with predictability, and the balance of randomness with determinism. The restriction of possibilities is important both on a higher level for the creative process, and on a lower level in terms of embodied interaction, or providing resistance. This is echoed in studies by Bertelsen et al. (2007), Magnusson and Mendieta (2007) and Armstrong (2006).

3.4 Capturing Embodied Interaction

To summarise this chapter so far, a design scope has been presented for forthcoming projects in this thesis, which is to create interfaces for interacting with GUIs that focus on higher

levels of embodied interaction than conventional input devices. This is to say, the interfaces will attempt to leverage human perceptual-motor skill in order to provide an improved quality of interaction with the computer.

An input device is typically a mediator from body motion to digital information. Typical GUI input devices such as the keyboard, the mouse and game controllers such as the joystick capture hand and wrist motion, finger motion and pressure. They do this with low dimensionality, and afford relatively small degrees of freedom of motion in interaction. For example, a mouse captures two continuous streams of motion information, and binary finger pressure readings from the buttons. A keyboard captures binary readings, and implicitly captures hand and finger motion as the hand moves from key to key. These interfaces both capture movement from the hand, and only exploit a small subset of the motion that the hand is capable of conveying. To leverage more of the hand's motion capabilities, or indeed any other bodily motion, an interface could be enhanced broadly in two ways: the amount of motion data which is being digitised can be increased, and the number of degrees of freedom afforded by the physical interface can be increased.

3.4.1 DEGREES OF FREEDOM OF MOTION

The degrees of freedom of motion afforded by an interface is tightly linked with its expressivity (Pressing, 1990). Take the example of the Theremin (Glinsky, 2005), which is considered to be a highly expressive instrument (Jordà, 2005, section 6.4.3). Like the computer mouse, it captures two continuous streams of motion data, although these parameters are operated independently, one with each hand. Unlike the mouse, the number of ways in which the body can be used to influence these two parameters is huge. The theremin is operated without physical contact, whereby the distance of the players' hands from two radio antennas controls the volume and pitch of a single oscillator. With no tangible interface, the player is free to use their hands in any way to influence the parameters. The degree of freedom of motion it affords is very large, and this has a significant effect on the expressiveness of the instrument, despite that only two parameters are being controlled. The same is true of acoustic instruments in general. Take the example of the djembe. The djembe is a drum played with the hands; the hands are used for striking and muting the drum head. Any hand shape can be used to create or modify the sound, so the number of degrees of freedom of motion is huge. Looking at the example of the violin as well, there are a multitude of possibilities for bow motion and finger motion.

In these examples, subtle changes in motion will tend to have an effect on the sound. This is not always so; some interfaces afford wide degrees of motion but they have a limited affect on the output. Let's consider some typical electronic music hardware. A slider, although outputting a single parameter, can be used virtuosically; a good example of this would be in turntablism. Due to the physical nature of the slider there are not nearly as many degrees of freedom as, for example, a hand controlling the pitch of a theremin. Subtle changes in control motion will make less difference to the outcome. Taking the example of a very unexpressive control element, a binary button, there are in fact many ways of pressing this (as can be seen at some computer music performances!) and many degrees of freedom of motion. However all

will produce exactly the same result. These examples show that when considering degrees of freedom of motion, it's more important to consider the *effective* degrees of freedom (EDFM), i.e. the number of possibilities of motion that can make a difference to the sonic outcome.

This concept links the physical nature of an interface with its control sensitivity. Increasing control sensitivity is one aim of the next concept to be considered in capturing embodied interaction: multiparametric control.

3.4.2 MULTIPARAMETRIC CONTROL

A multiparametric approach to capturing human motion for embodied control is an approach advocated by several studies. The basis of the multiparametric approach is simple; where many interfaces afford control through a sequence of single linear motions, a better approach is to capture several or many parameters at once, all controlled through a single motion. The multiparametric paradigm has long been implicit in the NIME field (Waiswicz, 1985; Pressing, 1990; Ryan, 1991), where designers have created new instruments using multiple sensors, mapping multiple data streams to collections of synthesis parameters. In the context of this project, this approach is extended to the design of interfaces for the control of digital music software tools.

Outside of the NIME field, Djajadiningrat et al. (2007) have begun to explore this area in the context of product design, and (see also section 2.5.1) propose that capturing control streams in parallel can increase the level of control while decreasing the number of actions needed for control. This approach allows the use of more complex movements such as twisting, rotating, squeezing, pushing, thereby enabling skilled control and expressive motion. They emphasise that the aim of this approach is not to create more complex action, but to allow for more sophisticated control.

Multiparametric control creates more complex data streams than the more conventional linear approach. One of the key issues concerns how to map this data to useful actions, and this is highly context dependent. In the context of musical interaction, mappings can be divided broadly into three groups: one-to-one, divergent and convergent (Rovan, Wanderley, Dubnov, and Depalle, 1997). With one-to-one mappings, one control parameter is mapped to a single musical parameter; this is the least expressive mapping. A divergent mapping describes a single control parameter that is connected with more than one musical parameter, and a convergent mapping is the opposite to this, where two or more control parameters influence a single musical parameter. Divergent mapping may limit detailed expression, as the single parameter will control a subset of the possible combinations of the musical parameters it is connected to. Conversely, a convergent mapping may be highly expressive, offering multiple ways of influencing a single musical parameter. Controlling a parameter in this manner will involve more skill than a one-to-one mapping; it may be hard to master but may also be more rewarding to control. The mappings present in acoustic instruments are in fact a complex system of both convergent and divergent mappings, with the additions of weights and thresholds (Hunt and Kirk, 2000). These kind of complex mappings are frequently found in the real world, but rarely in computer interfaces. Hunt and Kirk proposed that multiparametric control in this manner could encourage what they describe as *performance mode* interaction

(see section 2.3), and conducted a comparative study of input devices for musical control, one of which implemented this style of mapping. Their multiparametric system consisted of a combination of a computer mouse and two sliders. Different combinations of the slider positions and velocity, and mouse button state, position and velocity controlled the volume, pitch, timbre and panning of a sound, thus implementing convergent, one-to-one and divergent mappings. Users tended to be initially confused by this interface, but soon adjusted to the holistic manner of operation, focusing on whole gestures rather than individual parameters. Hunt and Kirk found that the multiparametric interface allowed spatial thinking, leveraging a natural human skill as part of the interaction in terms of embodied control. It also elicited subconscious control; users found it fun to use and felt it had long-term potential compared to the other non-multiparametric interfaces they tried. They concluded that non-linear mappings are more engaging for users, and that realtime control can be enhanced with a multiparametric approach.

3.4.3 A CONVERGENT MODEL OF INTERACTION

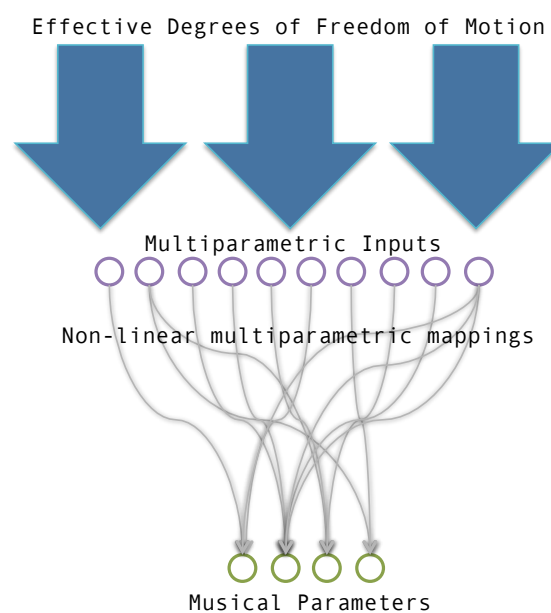


Figure 3.3: A Convergent Model of Interaction

It can be seen that these two factors, the effective degrees of freedom of motion and multiparametric data capture, are key factors in creating an interface that may afford a higher level of embodied control. Putting these two factors together, a convergent model emerges. This is illustrated in figure 3.3, which presents an idealised model of how motion could be captured in an embodied interface. The interface has a high EDFM, and then uses multiple inputs to capture motion data, mapping this non-linearly to musical parameters. The system gradually converges from a large number of control possibilities in the physical world, towards a much smaller number of musical parameters in the digital world. The high EDFM and large number of input parameters are properties of the physical interface. The mapping may be an inherent property of the physical interface, and may also be explicitly specified in the digital

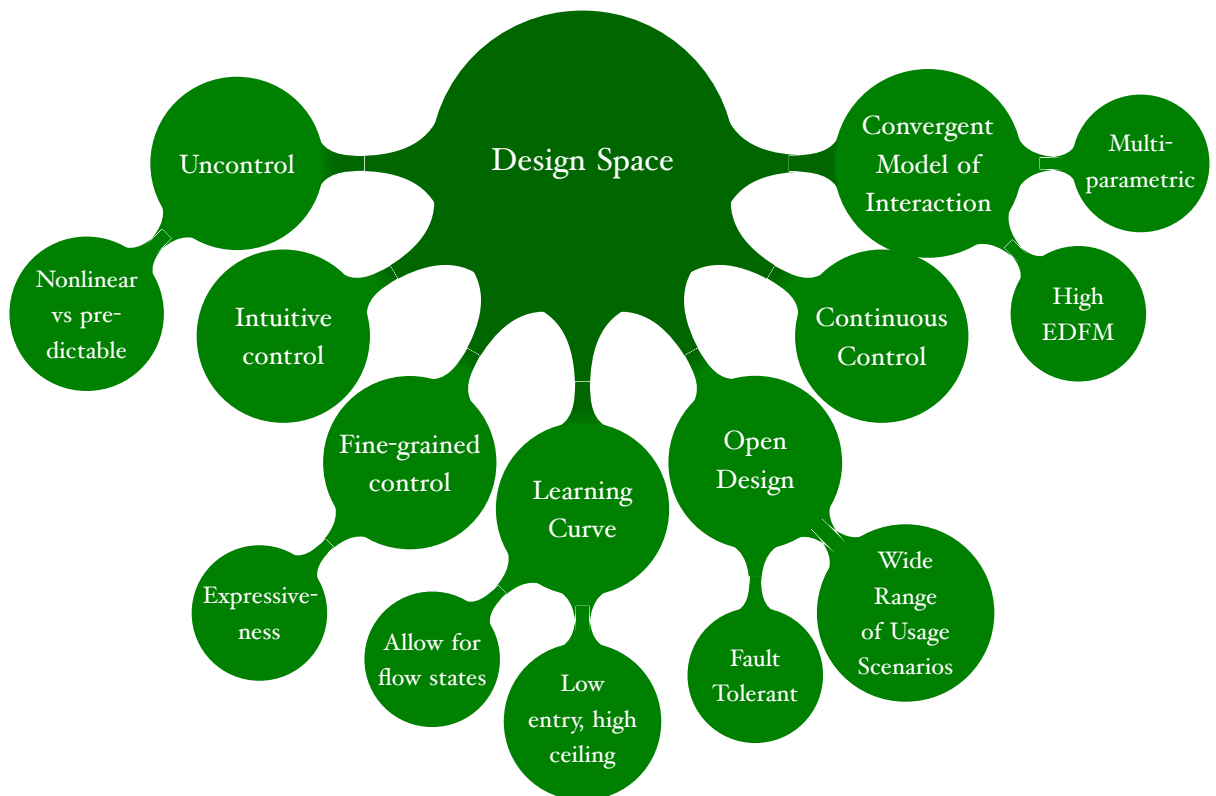


Figure 3.4: Design Space

domain. Based on the literature explored so far, an interface designed according to this model may be expressive, intuitive, and rewarding to use, but may also take time and skill to master. The efficacy of this model will be reviewed in the chapter 8.

3.5 A Design Space

At the beginning of the chapter, a design scope for new interfaces was presented that addresses the problems outlined in the previous chapter. The design scope focuses on input devices for personal computer interfaces, and ways of increasing the quality of embodied control of music software. A number of different and related design perspectives have been explored which are appropriate to this problem. From these, I would like to condense the salient features into a design space which will be used for the projects in this thesis. This design space is illustrated in figure 3.4. Examining this space in detail, there are four different types of control specified: uncontrol, intuitive control, fine-grained control and continuous control. Uncontrol is rooted in creativity and musicality; unpredictability is an important part of the creative process, and when balanced well with predictability, becomes an important part of musical interaction. Fine-grained control is also an important factor in musicality, allowing expressiveness and eventually virtuosity. Intuitive control and continuous control are both important factors in embodied interaction; intuitive control implies that an interface will build on our human perceptual-motor skills, and continuous control means that the condition of timely interaction can be fulfilled. Moving on from control, the convergent model of interaction

is an important part of the design space, a large factor for affording embodied interaction. Openness of design is also an important feature; this is the paradoxical aim of designing for unintended (mis)use. In practice, the designer can try to remove limits on potential usage scenarios, and try to build in fault tolerance. This sets up the potential for creative users to try the interface in different ways. Lastly, a low entry, high ceiling learning curve is important for setting up a path to virtuosity, and also for allowing flow states by balancing the challenge and skill needed for using an interface.

This design space will be used later in this thesis as a framework within which to place the evaluation results from the systems I have developed.

3.6 Emergence and Real-world Constraints

Having established a design space, the real-world constraints on the projects presented in this thesis must also be examined. While this design space represents an ideal to aim for, there are other pragmatic factors involved in the creation of the interfaces which must be acknowledged. These constraints have been the availability of materials, technology, facilities and software libraries. Overall, I have taken a bottom-up approach to these projects where I have tried to put together these interfaces as much as possible from proven building blocks. This has been done with the objective of trying to create working prototypes in a reasonable time-frame without getting slowed down in low-level development of software or hardware; of course low-level work has been involved, but neither of the two interfaces have been made completely from scratch, all make use of established hardware or code libraries on some level, most of which has been fine-tuned in some way to suit the project. In general, the design process has started with a wider goal in mind, along with the aims from the design space. From here, I have experimented with different hardware interfaces, and then tried a variety of software mapping techniques from different libraries, in a variety of scenarios. From these scenarios, the system has been fine-tuned to reach a state where evaluation can start. These interfaces have also been developed with a low hardware budget, which has been an inspiring constraint and also fundamental for making low-cost interfaces which can be reproduced and sent out for field evaluation. Overall, the process has been very explorative, and I wish to emphasise this because it sets the storyline over the next group of chapters. The two interfaces were created through an emergent process which was influenced and inspired not only by the higher design goals but also the environment surrounding their creation.

3.7 Summary

A design scope was established, based on the research aims outlined in the previous chapter. Following this, design guidelines from various fields have been explored, from creativity support, flow, digital lutherie and studies of computer music composers. A convergent model of interaction has been proposed, with the aim of creating interfaces which realise embodied control, and capture detailed motion. Putting this together, a design space was introduced,

based on the salient features from the design guidelines that have been explored.

The complementary process to design is evaluation. This evaluation of musical controllers presents some difficult and interesting challenges, and these are explored in the next chapter.

CHAPTER 4

HCI METHODOLOGY FOR EVALUATING MUSICAL INTERFACES

‘If you want to penetrate the mind of an
artist, you must visit him in his studio.’

Robert Schumann, 1856. (Landis, 2000)

4.1 Introduction

The conventional interaction design process involves the iterative creation of specific design goals, their implementation and then evaluation (Preece, Rogers, and Sharp, 2002). In the specific context of music, the nature of the way we interact with tools creates obstacles to the use of conventional evaluation techniques. This chapter explores the relationship between evaluation and the design of creative tools, in order to arrive at a set of research methods that will be used to evaluate, as best as possible, the work carried out in this thesis.

4.1.1 THE CHALLENGES OF EVALUATING MUSICAL CONTROLLERS

There is no such thing as a perfect creative tool. This may seem an obvious statement, but it's worth digging a little deeper into this to explore some issues which are relevant in this chapter. Take the example of Miller Puckette, creator of the software Max/MSP. Discussing the design of this system, he acknowledges that the designer cannot predict all aspects of intended use (*Even though good software writers can themselves dream of many things, the software user can always think of something else. Although we software designers try above all else to avoid imposing restrictions or obstacles, we never succeed entirely*) (Puckette, 2002). Further to this, he attributes the early success of the software to factors that include the creative abuse of the system by its users (Puckette, 1996). In saying this, he is pointing out a fundamental paradox in the design of creative tools; real-world use will always move beyond the scope of the design, and the success of a system partly depends on this out of scope use, yet how can you design for scenarios you can't predict?

Magnusson and Mendieta's survey of digital musicians explores *creative abuse* from the player's perspective; they found that the limitations of an instrument were a source of inspiration, and that part of their creative process was exploration of these limitations. Players liked to find the boundaries of an instrument and push against them (Magnusson and Mendieta, 2007). Bertelsen et al. (2007) encapsulate this pushing of boundaries in the concept of *materiality*. They describe instruments as having material resistance; this describes not just physical resistance, but also conceptual or virtual resistance in software in the form of text, interface element, metaphors and algorithmic behaviours. In their case study of two composers, they also discovered this process of pushing against limitations as part of the creative cycle. Like Magnusson and Mendiata's respondents, the two composers found these limitations inspiring. Bertelsen et al. suggest that materiality is what makes the software an instrument, it provides embodied resistance for the musicians to play on. These two examples highlight the role of limitations and resistance in a system, and demonstrate further the creative tool design paradox.

Coming back to the original statement, that there is no perfect creative tool, these examples demonstrate that a lack of perfection is a fundamental part of a creative tool. From the players perspective, imperfection can be creatively inspiring, and the probing of limitations is part of the creative process. Designers need to acknowledge this, and also accept that their creation is going to be used in ways beyond where they can envisage. This situation creates some complex challenges for evaluation. With such open design goals, where should

evaluation be focused? And what is the best manner to approach the process?

4.2 Evaluating the Wiimote

In order to experiment with different techniques and gain a deeper understanding of the evaluation process, a pilot evaluation study was carried out. This study was an evaluation of Nintendo's Wiimote as a musical controller. The full details of this study will be presented to provide a context for later discussion.

At the time of this study in 2007, Nintendo's Wii gaming console was enjoying considerable success; 20 million were sold worldwide, more than any of its rivals¹. The reason for this popularity could be partially attributed to the innovative design of its controllers, whose motion sensing capabilities introduce gestural control into gameplay. The console's principal controller, nicknamed the Wiimote, can function independently from the Wii. It sends data wirelessly using the Bluetooth protocol, which means that ordinary computers can read its output. This output can in turn be directed to audio software, enabling the device to be employed as a musical controller.

The Wiimote has been a device of particular fascination for the electronics hacking community, and has been appropriated for all manner of HCI projects (Chung, 2008). The scale of musicians' interest in the Wiimote can be observed from the proliferation of demo videos and community sites on the internet. The Wiimote was the first affordable and commercially available accelerometer based controller, and the style of interaction it affords struck a chord with computer musicians. People found applications for the controller which include drumming (Burkhart and Moebert, 2007), conducting (Bruegge, Teschner, Lachenmaier, Fenzl, Schmidt, and Bierbaum, 2007), DJing (DJ WiiJ, 2010), synthesiser control (Buskirk, 2006) and more. This interest was reflected by computer music software developers, with Wiimote extensions available for many audio environments. Generic solutions such as GlovePIE² or DarwiinRemote OSC³ allow the conversion of Wiimote data to MIDI or OSC.

The device was in use by musicians and yet was not specifically designed for this purpose. This meant there was an opportunity to investigate the Wiimote in an attempt to gain a better understanding of its capabilities in a computer music context. To this end, and also as a vehicle for the exploration of evaluation techniques, a study of the Wiimote as a music controller was carried out. Human-computer interaction methodologies provided the basis for conducting the study. As well as wider HCI literature (Dix, Finlay, Abowd, and Beale, 2004; Jacko and Sears, 2003), the design of the study drew on research into evaluating musical controllers by Wanderley and Orio (2002), and Höök, Sengers, and Andersson's (2003) study on evaluating usability in interactive art.



Figure 4.1: The Nintendo Wiimote

4.2.1 THE WIIMOTE

To determine how to evaluate the Wiimote, the musical possibilities afforded by its control set need to be examined. The Wiimote embodies three types of control: a three-axis accelerometer for motion sensing, an infrared camera for pointing, and various buttons. For the purposes of the study the accelerometer was the main focus, at the time a rare feature in musical controllers and arguably the most interesting feature for musicians.

4.2.1.1 THE ACCELEROMETER

The effect of gravity on an accelerometer means that it can be used to measure rotation about the Z (roll) and X (tilt) axes, though yaw cannot be measured as gravity has no bearing on the device when rotated about the Y axis (Mizell, 2003). These roll and tilt readings can be used for continuous control, with two caveats; firstly, accuracy only comes from pure rotation, additional lateral motion causes acceleration that will add noise to the data. Secondly, the output is not precisely linear, approximating instead to a slight 'S' shape with values bunching up in the centre.

The raw acceleration data can be interpreted in different ways. One possibility is for triggering; a peak detection algorithm can determine when a drumming-like motion has been made, and in turn trigger an event such as playing a percussion sample. Another possibility is gesture recognition, which presents several challenges. Acceleration data from the Wiimote

¹Console wars. http://en.wikipedia.org/wiki/Console_wars

²<http://carl.kenner.googlepages.com/glovepie>

³<http://code.google.com/p/darwiinosc/>

is inherently noisy; if it is rolled or tilted then gravity affects the readings and it is difficult to remove this component to get the true acceleration value. Conversely, a rotation may look like an acceleration movement. For this reason, it is difficult to determine absolute position reliably, so gestures must be inferred somehow from the raw acceleration values. Neural networks (Bishop, 1995) are well suited to solving this sort of classification task. Having examined the basic capabilities of the Wiimote for musical control, a study was planned to test them.

4.2.2 METHODOLOGY

The field of Human Computer Interaction provides tools and methodologies for evaluating computer interfaces (Dix et al., 2004), but applying these to the specific area of computer music can be problematic (Jordà et al., 2007). HCI methodologies have evolved around a task-based paradigm and the stimulus-response interaction model of WIMP systems, as opposed to the richer and more complex interactions that occur between musicians and machines. Höök et al. (2003), discussing the relationship between HCI and installation art, suggests that art and HCI are not easily combined, and this is also true in the multi-disciplinary field of computer music.

At the time this study was run, at the end of 2007, there was limited HCI literature which focussed specifically on computer music and evaluation. Since the study there have been more contributions to the field, notably Stowell et al's work on discourse analysis (Stowell, Robertson, Bryan-Kinns, and Plumbley, 2009) and Gelineck and Serafin's (2009a) evaluation of knobs and sliders. This study however, took place in the context of the research now discussed.

Höök et al. (2003) examine the use of HCI in interactive art, an area which shares common ground with computer music. She describes her methodology for evaluating interaction in an installation, and examines the issue of assessing usability when artists might want to build systems for unique rather than 'normal' users; music shares similar characteristics with art. Poepel (2004) presents a method for evaluating instruments through the measurement of musical expressivity. This technique is based on psychology research on cues for musical expression; it evaluates players' estimations of a controller's capability for creating these cues. Wanderley and Orio (2002) have conducted the most comprehensive review of HCI usability methodologies which can be applied to the evaluation of musical systems. They discuss the importance of testing within well defined contexts or metaphors, and suggest some that are commonly found in computer music. They propose the use of simplistic musical tasks for evaluation, and highlight features of controllers which are most relevant in usability testing: learnability, explorability, feature controllability and timing controllability. Their research fitted best with objectives for evaluating the Wiimote, and had the largest influence on the design of the study.

4.2.3 THE STUDY

4.2.3.1 METAPHORS AND MUSICAL TASKS

Following guidelines from Wanderley and Orio, the study comprised asking participants to perform simplistic musical tasks within metaphors which were chosen to test the basic capa-



Figure 4.2: The HandSonic and the Wiimote Used in the Study

bilities of the Wiimote. The study was a comparative study, asking participants to perform these musical tasks on an additional controller other than the Wiimote. The Roland HPD-15 HandSonic⁴, pictured in figure 4.2, was chosen for this purpose as it provided the interface elements that might typically be used for performing these equivalent tasks. The metaphors and musical tasks shall now be described, and the data collected for each:

Triggering

Participants were asked to trigger drum samples by making drumming-like motions with the Wiimote. They played simple patterns in time with a metronome beginning with crotchets, moving up to quavers and then any pattern they chose. They performed the same tasks using the HandSonic's drum pads. Drum trigger events and metronome beats were logged for later analysis.

Precise Continuous Control

Continuous input controlled the pitch of a saw wave in 6 discrete steps. Participants were asked to move up and down through successive pitches in time to a metronome, using both the roll and tilt axes of the Wiimote and also a knob on the HandSonic. The pitch changes and metronome beats were logged.

Expressive Continuous Control

This task involved simultaneous control of the grain density and filter parameters of a generated sound. These parameters were mapped to the roll and tilt axes of the Wiimote, and to two knobs on the HandSonic. Participants experimented with each controller for about two minutes. The less defined nature of this and the next context would not lend themselves well to statistical analysis; no quantitative data was collected for either.

Gesture Recognition

A multilayer neural network was trained using back propagation to recognise 5 different

⁴http://www.roland.co.uk/drum_room_catdet.asp?ID=HPD15

shapes, which could be drawn while holding down the B button on the Wiimote. Five rhythmic musical tracks were running simultaneously; each shape was assigned to a track and recognition of the shape muted or un-muted this track, quantised to the nearest bar. Participants were given a printout of the gestures and asked to play with the system for 3-4 minutes. There was no comparison controller in this part of the study.

4.2.3.2 PARTICIPANTS

There were 17 participants, with an average age of 31.2 (min: 20; max: 46). For the purpose of statistical analysis, nine were classified as musicians based on a combination of years of study and practice routine. They had either completed at least six years of training, or at least two years and were practising at least six hours per week. The musicians averaged 7.7 years of study (min: 2; max: 15) and 6.8 hours of practice per week (min: 0; max: 25). Participants also rated their Wiimote experience on a scale of zero (no experience) to five (expert); most participants had no prior Wiimote experience; wiimote-gamers were classified as having two or more for this value, giving six wiimote-gamers with an average experience of 3.2.

4.2.3.3 METHOD AND IMPLEMENTATION

Each session of the study started with the participant being asked about their musical and Wiimote experience. Before each context they were given a period of practice with each controller, then afterwards they were interviewed about their experience. The order of the controllers and Wiimote axes was randomised between participants to reduce learning effect. After the first three contexts they were questioned about which controller they preferred and why, and asked to describe the advantages and disadvantages they felt the Wiimote possessed in that particular metaphor. With no comparative controller, the interview after the gesture recognition context was more open ended; participants were asked to comment on their experience and were questioned on certain aspects of the task. Having completed the tasks, they were asked some general questions about their experience.

4.2.3.4 IMPLEMENTATION

The software for the study was programmed using the SuperCollider audio programming environment (McCartney, 2002), which was connected to the Wiimote via DarwiinRemote OSC and to the HandSonic via MIDI. This software also recorded data logs. Participants were videoed in order to observe how they used the Wiimote and to record their answers to the interview questions. The logged data was analysed in MATLAB (as later described), and a qualitative analysis of the video interview data was conducted, identifying and coding common themes.

4.2.4 RESULTS

4.2.4.1 QUANTITATIVE RESULTS

Data was analysed using ANOVA and post-hoc t-tests with respect to factors of Wiimote/HandSonic, musician/non-musician and wiimote-gamer/non-wiimote-gamer as appropriate.

Table 4.1 shows the participants' preferred controllers for the first three tasks. There was no significant overall preference for either controller.

For the triggering test, the timing of each successful trigger was logged in the crotchet and quaver tasks, and the data analysed to determine the average timing error relative to the task requirement. ANOVA tests revealed no significant difference between the Wiimote and the HandSonic, both overall and between sub-groups of musician/non-musician and wiimote-gamer/non-wiimote-gamer.

Pitch changes during the precise continuous control phase were analysed to obtain the frequency of changes and timing errors for each of the roll axis, the tilt axis and the knob on the HandSonic. There was a borderline non-significance between the number of pitch changes from the roll and tilt axes ($p=0.0542$), pointing towards a tendency for the tilt axis to be more accurate for control. Comparing the roll and tilt axes to the knob via individual t -tests gave highly significant differences ($p<0.00001$, $p=0.0003$ respectively), which was expected considering the stability of a knob compared to the Wiimote. No significant difference was found in timing error between the three control methods.

	Wiimote	HandSonic	Neither
Triggering	18%	70%	12%
Precise Cont. Ctrl	35%	53%	12%
Expressive Cont. Ctrl	53%	23.5%	23.5%

Table 4.1: Controller Preferences

4.2.4.2 INTERVIEW RESULTS

The interviews were transcribed and analysed using a grounded theory approach (Strauss, 1987; Strauss and Corbin, 1998). The results shall now be summarised, presenting the themes that emerged and the participants' comments about them.

Triggering

The main focus was on the lack of physical feedback with the Wiimote. The absence of a real contact point made the task difficult for some; one participant solved this by drumming against their hand. Some people commented on the intuitive nature of the Wiimote during this task (*it has the feeling of a virtual drum stick*), and many talked about the benefits of the controllers portability. It was observed that some participants found difficulty in drumming faster rhythms, especially semiquaver patterns.

Precise Continuous Control

Issues emerged mainly in the areas of control, mapping, ergonomics and feedback. People commented that they found it easy to get stuck between notes, that was it difficult to judge the boundaries, and that the Wiimote was generally less steady and precise than the HandSonic's knob. On the positive side, some liked the speed and freedom of movement compared to a knob. Several participants noticed how the control wasn't precisely linear. Ergonomically, some people found the 360° rotation action unnatural for their wrist. Participants talked about the lack of visual feedback with the Wiimote, preferring the HandSonic where they had a viewable reference for the controller setting.

Expressive Continuous Control

Fun was a prominent theme here, some people feeling that the Wiimote was a more

enjoyable way of controlling the sound, especially as precision wasn't required. Many participants talked about the intuitive or embodied feel of using the Wiimote in this context with comments such as *'it's almost like your own hand making the noise'*, *'I had to think a lot more using the HandSonic'* and *'it's more instrument like, less computer like'*. Expressiveness and musicality was a common topic; some felt it gave more possibilities (*'you can explore stranger noise combinations'*) and more room for expression. People observed that you could control the parameters percussively, and some enjoyed the randomness the Wiimote added to the control process. In terms of control, there were comments about repeatability (*'it's hard to stay in one place, which could be good or bad depending on context'*) and co-dependence (*'it's easy to affect one parameter while changing the other'*). Several people said they appreciated controlling two parameters with one hand. Physical feedback was mentioned again (*'I prefer the hard limits of the HandSonic'*).

Gesture Recognition

Participants again emphasised the fun aspect of using the controller (*'it's fun way of turning stuff on and off'*, *'keep fit with drum loops!'*). Opinions were mixed on the subjects of intuitiveness (*'it felt like a biological relationship between me and the music'*, *'it was strange to control with shapes, I didn't really feel part of it'*) and expressiveness (*'didn't feel more expressive than playing a button'*, *'making a shape feels more dramatic and connected to the music'*). One participant observed that this task was the most analogous to Wii gaming. In terms of gestures, several people would have preferred to make smaller more subtle gestures, and some said they would have liked some sort of continuous measurement of the gestures, rather than just binary control. One participant would have liked less arbitrary gestures that related more to the sound. People thought that this kind of application of the Wiimote would be good for performance, and would look good on stage.

General Comments

Participants were asked about what expectations, if any, they had of the Wiimote in terms of music. The majority of people said they had none, and two said they thought it would have been less responsive. When asked whether they imagined any sort of metaphors while using the Wiimote, most participants said they didn't. Some people imagined the Wiimote as a drum stick in the first task, although this was problematic for one person (*'it makes you want to use it like a drum but it's not like a drum. If you think of it as a drum stick, it's difficult to get it to do what you want'*). One person thought of rolling the controller as being like turning a knob.

About whether they could imagine using the Wiimote in their own projects, several people thought it would make a better tool for performance than composition (*'I'd use it to make performance more of a spectacle'*). Guitarists talked of attaching it to their guitars for an extra dimension of control. A singer proposed using it for controlling vocal effects during performance. Other people suggested strapping it to limbs or using it for conducting.

Participants were also asked how they thought the Wiimote could be improved, either physically or functionally. A strong theme was absolute positioning, participants believing that this would make the device more useful. Another theme was virtuality, participants wanting the device to have some sort of physical feedback or visual feedback such as a laser pointer at the end. The shape of the device was an issue (*'it's like a TV remote control'*), with suggestions for a malleable surface or a rubber grip. Someone commented on the weighting (*'weighting and balance is important, different sounds need different weightings'*). In terms of additional controls, there were suggestions that larger buttons would be better for music and that some sliders would be useful.

4.2.5 DISCUSSION

This study was set up to explore more systematically how the Wiimote functioned as a music controller, comparing it to another type of controller both in terms of actual performance and user experience. While the statistical analysis showed no significant differences, there were some interesting themes arising from the interview data about the experience of using the Wiimote.

Two overall themes emerged from the interview data: virtuality and expression. The abstract nature of interaction with the Wiimote seemed to be the main contributor to control issues that some participants had with the device: lack of hard limits for continuous control, lack of physical feedback, lack of visual reference and lack of a concrete metaphor. One participant commented on how this abstraction made the triggering task difficult: *‘the virtual nature of the Wiimote makes it harder to keep the rhythm in my head, after a while I started to lose the meaning of what I was doing’*. This virtuality is also one of the Wiimote’s strengths, giving it flexibility for use in multiple contexts and providing one of the roots of the embodied and intuitive experience some participants observed during the study.

The expressive continuous control context was the only metaphor where the Wiimote won the title of *most preferred controller*, and this points to where some of the Wiimote’s strengths lie for musicians. The physical nature of accelerometer control lends a natural instability to its output, and the addition of acceleration motions such as ‘flicking’ the controller add an extra and slightly unpredictable dimension to the possibilities. This manner of less precise control sits very comfortably in the context of creative or expressive sound manipulation. This is not to say that the Wiimote isn’t useful in more structured contexts. From a functional point of view, there was no overall difference in timing error between the two controllers for the time critical tasks, so the Wiimote was just as capable as the HandSonic for timed accuracy at the resolutions measured (quarter and eighth notes).

4.2.6 SUMMARY

The potential musical applications of the Wiimote were examined, and its usability was evaluated in what was believed to be some of the core contexts of its use. This evaluation shed light on some of the problems that might occur when employing the device, and on situations, such as expressive control, where the controller may yield more creative potential. The Wiimote is relatively cheap and easily connectable to home computers, making it widely accessible to musicians. The results show it can add interesting and novel dimensions to musical control, provided that some limitations are accounted for.

4.2.7 REFLECTING ON THE WIIMOTE STUDY

Having presented the implementation and results of the Wiimote evaluation, it is also useful to reflect more generally on the structuring of the study and the efficacy of the HCI evaluation. Was it useful to carry out the Wiimote usability study with the methods chosen? Where were the gaps in the results and how could the methodology be improved to narrow these gaps?

Focusing on the data, the most interesting and useful results came from analysis of the

interview data. The interviews confirmed some expected results about the controller but more usefully brought up some unexpected issues that some people found with certain tasks, and some surprising suggestions about how the controller could be used. This is the kind of data that shows the benefits of conducting a usability study, the kind of data that is difficult to determine purely by intuition alone and that is best collected from the observations of a larger group of people. From the remaining results, the quantitative results provided objective backup to certain elements of the interview results, some useful data about the functional side of the controller, and insight into global trends of the participants. However, the conclusions reached from these results alone seemed to be a limited measure of the device compared to the subtlety and detail of the participants' observations.

Did the study result in a complete answer in relation to the research question, *how useful is the Wiimote as a musical controller?* It's difficult to answer this objectively, but it can be observed that the results showed a detailed and intimate understanding of the controller in a musical context.

There are two areas where the picture painted by the Wiimote study is incomplete. First, the very nature of running a laboratory based study means that the study did not evaluate real-world usage. It evaluated very specific features of the controller in a controlled environment, and this in turn elicited more general reactions to the controller from the participants. While the lab-based study has the advantages of controlling experimental variables, it is very far removed from how an instrument is used in reality, and this must colour the results. The other missing element was any objective measurement of the participants' experience in the moment of using the controller. More interesting results came from post-task interviews, but there is no data about experience in the moment while using the device, something that would seem important for a musical evaluation. This gap in the results is partly due to lack of technology and partly due to a lack of standardised methodology. How can musicians self-report their experience while they are using a musical controller without disrupting the experience itself? Are there post-task evaluation techniques that can give a more accurate and objective analysis of a musical experience than an interview? More recent research in HCI is starting to address similar issues and can point to possibilities.

4.3 Evaluation Methodology

4.3.1 FROM USABILITY TO USER EXPERIENCE

Kaye, Boehner, Laaksolahti, and Staahl (2007), in 2007, described a growing trend in HCI research towards *experience focused* rather than *task focused* HCI. With this trend comes the requirement for new evaluation techniques to respond to the new kinds of data being gathered. This trend is a response to the evolving ways in which technology is utilised as computing becomes increasingly embedded in daily life, a shift in focus away from productivity environments (Mandryk, 2005), and from evaluation of efficiency to evaluation of affective qualities (Fallman and Waterworth, 2005). As HCI is increasingly involved in other 'highly interactive' fields of computing such as gaming, the requirement for evaluating user experience becomes

stronger. This new trend is known as the ‘third paradigm’, and researchers have started to tackle some of the challenges presented by this approach. Some of these techniques will now be explored, starting with physiological measurement.

4.3.2 PHYSIOLOGICAL MEASUREMENTS

For highly interactive tasks such as playing a musical controller, a non-interactive data gathering mechanism is essential to gain realtime data, so the measurement of physiological data (Fairclough, 2008) may be a promising technique for yielding realtime readings without interrupting the users’ attention. Several studies have focused on this area. The AMUSE system (Chateau and Mersiol, 2005) is designed to collect and synchronise multiple sources of physiological data to measure a user’s instantaneous reaction while they interact with a computer system. This data might include eye gaze, speech, gestures and physiological readings such as EMG, ECG, EEG, skin conductance and pulse. Mandryk (2005) examines the issues associated with the evaluation of affect using these physiological measures; how to calibrate the sensors and how to correlate multi-point sensor data streams with single point subjective data. Both studies acknowledge that physiological readings are more valuable when combined with qualitative data. Practically, these techniques are not simple to apply. Numerous factors, both personal and environmental, may influence readings, and care needs to be taken to obtain true measurements (Mandryk, 2008). Furthermore, readings vary from individual to individual and from day to day, making global comparisons difficult. Mandryk’s work is based in the field of computer gaming (Mandryk, Inkpen, and Calvert, 2006), where this research is becoming more widely applied. Game designers are evaluating player experience in order to find out where their games fall short in terms of excitement and challenge, and biometric measurements can help to pinpoint these areas (Hazlett, 2007; Mirza-Babaei and McAllister, 2010a,b). Although there are similarities in interaction between gaming and music, the application of these techniques for computer music evaluation may prove problematic, and as yet little work has been done in this area. A major stumbling point could be the difficulty in separating out the player’s affective reaction to the instrument from the player’s affective reaction to the music they are playing. With the lack of research into the specific application of these techniques to computer music interface evaluation, these physiological approaches have not been attempted for the studies described in this thesis.

4.3.3 REAL-WORLD MUSICAL INTERACTION - EVALUATION IN THE WILD

To understand the importance of evaluating real-world usage, we need to step back a little and examine how musical interaction takes place. Looking closely, it becomes clear that to obtain a full picture of a controller in use, research must take place outside the lab, and must also take place over enough time for the musicians to become familiar with and skilled with an instrument. As we have seen earlier, creative use of an instrument can involve usage outside the designers intentions or expectations. These kind of scenarios emerge with natural usage, as the player becomes acquainted with a new system, finding its affordances, resistances and limitations. Musical interaction in itself is explorative, involving an emerging relationship between player and instrument that evolves over time. This is particularly the case with

computer music systems, which are rarely learnt with a teacher, but more commonly picked up through exploration and improvisation by the player (Jordà, 2005). It's this process of exploration which leads to insights into the instrument and creative ideas for its use, and this is highlighted in several studies of electronic music composers. Studies by Bertelsen et al. (2007) and by Gelineck and Serafin (2009b) both observe how probing an instrument is part of the creative process. Gelineck and Serafin comment that *'Participants seem to prefer working in a free exploratory mode early on in the compositional process. They explore new ideas by trying to break boundaries, interact with musical tools differently than intended'*.

Many aspects of musical interaction are only meaningful in specific social contexts, for example public performances, studio work, rehearsals, informal collaborations, tuition sessions and more. Each of these scenarios exposes a different facet of interaction between player and instrument. Considering public performance as an example, important aspects of an instrument being tested will include repeatability, reliability and durability. In a different context, for example informal group improvisation, the explorability and shared interface elements may have greater importance; in solo practice sessions, aspects of learnability will be tested.

To obtain a true picture of musical interaction when evaluating a musical system, typical use needs to be allowed to take place. This is something that needs to evolve over time as the player discovers and tests the instrument, and it's socially situated, something that you can't replicate in a laboratory study. To this end, in order to fully evaluate a musical controller, a significant part of the evaluation should take place in the field, and therefore field studies were used as one part of the evaluation strategy for the studies in this thesis. This view is echoed in the wider field of creativity research. Shneiderman et al.'s (2006) guidelines for evaluating creativity support tools state that .

'Although controlled experimentation has long been seen as the leading approach to rigorous research in many areas of psychology and HCI, there was little sympathy for such methods in our workshop discussions. Controlled studies in laboratory conditions with standard or "toy" problems over a few hours were seen as inadequate to capture the strategy changes, new possibilities, and learning effects with powerful software tools, as they are applied to complex problems. More sympathy was expressed for in-depth longitudinal case studies and ethnographic field study methods to capture the rich texture of activity among creative individuals or groups.'

Having mentioned this importance of situated interaction in social contexts, at this point it is necessary to clarify the scope of this work with regards to social and collaborative music making. While there is some compelling research in the NIME field that explores musical interaction through the lens of collaborative music making (Bryan-Kinns and Hamilton, 2009; Jordà et al., 2007) and group audience participation (Sheridan and Bryan-Kinns, 2008), it is not the intention of this thesis to explicitly explore this area. In consistency with the theme of interfaces for digital music tools, the controllers were designed principally for solo use, and it was left to the field evaluations to see in which social contexts, if any, they would be used in.

4.3.3.1 DATA COLLECTION

Given that observing real-world usage is an important part of evaluating an instrument, what is a good way to collect real-world data? The requirements of observing musical activity present a set of constraints to the researcher. First, the typical nature of creative activity rules out participant observation in the field, which would be both impractical and interruptive. Musical activity is also not limited to a single location, so this rules out the use of situated observation equipment. This means that the data must be collected by the participants themselves; they must record their own experience. The method of recording data cannot interrupt the experience of using an instrument, so this data collection method will either reflect the participant's experience after the event, or will consist of some kind of recording initiated by the participant. Musical experience is multi-faceted, so useful data could be in the form of any media e.g. audio, video or text. Given these restrictions, in the field studies described in this thesis, participants were asked to collect data in a freeform way which would suit whatever their creative activity was. The data could take the form of a diary, which could be filled out after studio sessions, or media recordings, such as videos of the participant playing the instrument or sound recordings. The participants were also interviewed at the end of the period of the field study. At this final interview, media that had been collected by them was examined together with the researcher. In the same way that video is used with stimulated recall debriefing (Bentley, Johnston, and von Baggo, 2005), these artefacts helped the participants to connect with the memory of their experiences of using the controller more intimately.

4.3.4 COMBINING RESEARCH METHODS

Having professed the value of field studies for evaluating musical systems, it must not be forgotten that controlled lab-based studies can also be extremely useful. This is especially true in the early stages of development of a system where quick feedback is needed and potential areas of improvement need to be identified. These early studies are known as formative evaluation (Preece et al., 2002). The projects presented in this thesis have all taken a combined approach, of early formative evaluation of a system followed by field evaluation.

The formative evaluations have followed the template of the Wiimote study; the participants have been asked to perform specific tasks with the controllers in a laboratory-based setting, and have been interviewed to obtain their responses. A decision was made to focus on qualitative interview data rather than quantitative measurements. As demonstrated in the Wiimote study, this can give far more detailed and subtle information about a controller than statistics can. Further to this, at the formative stage of evaluation the design is still open and being explored conceptually rather than functionally. More value can potentially be obtained by opening up the field of enquiry with semi-structured interviews, rather than narrowing it by testing specific features whose importance in the wider picture is as yet unknown. Another factor in collecting qualitative data has been the scale of experimentation; to obtain enough data to be useful for statistical analysis, a relatively large number of participants is required such as in the Wiimote study. At a formative stage, conducting a large study is not always practical or desirable, and it can be more useful to obtain feedback quickly in a small study. This is not to discount quantitative methods in the evaluation of computer music systems,

but in the context of this thesis, qualitative data has been more appropriate.

Overall, the combination of these methods complement each other well. A lab-based formative evaluation can outline issues that need to be dealt with before a field study is carried out, and a field study covers areas impossible to include in a lab-based study, and can compensate for the methodological shortcomings of a controlled study. Combining the results helps to give a more complete picture of the controller than either type of study can independently.

Field studies have rarely been carried out within the area of computer music controllers, so one of the contributions of this thesis is to describe and evaluate the use of this methodology.

4.3.5 DATA ANALYSIS

Both formative and field studies collect data in the form of transcribed interviews, while field studies also collect other types of media appropriate to the experience of the participant. To analyse this collected data, a grounded theory approach was followed. Grounded theory (Strauss and Corbin, 1998; Strauss, 1987) is particularly appropriate because its primary purpose is discovery (Strauss and Corbin, 1998), which is well matched to exploring the reaction of musicians to new controllers and new styles of interaction. Grounded theory also fits in well with field studies because it can include multiple types of data in the same analysis, allowing for the combination of video, audio, and different text sources.

4.3.5.1 GROUNDED THEORY

Grounded theory is a style of analysis, developed originally for social science studies in the late 1960s by Glasser and Strauss (Glasser and Strauss, 1967). It provides guidelines rather than rules for the systematic analysis of data in order to generate a theory and test it. Where most research methods work by testing a preconceived hypothesis, grounded theory works by generating a hypothesis. This is achieved through the triad of data collection, coding and memoing. Coding describes the conceptual labelling and organisation of data, and memoing describes the writing of an analysis of data, which can take place on a microscopic level. The researcher develops an intimate relationship with the data, and is aware of themselves as an instrument in the process of the analysis. While this may pose problems with objectivity, Strauss and Corbin believe that *'the researcher's [interpretation] will not be the only possible interpretation of the data, but will be plausible, useful and allow its own further interpretations'*. The aim of grounded theory is to identify process and mechanisms surrounding a phenomena, contributing to the wider picture rather than defining it.

Grounded theory has been successfully used for analysis in a number of art related studies, for example Eaglestone's survey of electroacoustic composers (Eaglestone et al., 2008), and Mace and Ward's study of the creative process in art making. Mace and Ward appreciated the emergent nature of grounded theory analysis in relation to the subject of their investigation:

'Given the incomplete state of knowledge concerning what actually happens during the development of a work of art, we believe it is appropriate to let the artist's relatively unstructured descriptions of what he or she did inform us, albeit by the application of a systematic method, as to the creative process.' (Mace and Ward, 2002)

Grounded theory has also been used for evaluation of interactive art; Hohl interviewed visitors to his *Radiomap* installation in order to get a better understanding of their experience. He gives his reflections on methodology:

‘Although time consuming, it has been very rewarding to develop a Grounded Theory of our participant’s experience. By applying this qualitative method, we have gained a much better understanding of the manner in which participants perceive the interaction with the application in their own voices. This gave us detailed insight into the process of interaction itself, a knowledge resulting in unexpected findings we could not have determined by observation or interviews alone.’ (Hohl, 2009)

This very much echoes my own reflections on the qualitative analysis in the Wiimote study; the detailed analysis involved in grounded theory can provide subtle and detailed insights into participants’ experience, and is an extremely useful tool for the evaluation of musical interaction.

4.3.5.2 GROUNDED THEORY PROCESS

There are two schools of grounded theory, that diverged as the two original developers of the technique published their own separate texts. Investigations by Kendall (1999) and Heath and Cowley (2004) both conclude there is no correct answer as to which approach to follow; both have their own specific strengths and weaknesses, and both represent the same core process. In this thesis, Strauss and Corbin’s approach has been followed (Strauss and Corbin, 1998).

To illustrate the application of grounded theory in this thesis, the process of analysis will now be outlined. The majority of data was collected from interviews, which were recorded as audio. After interview sessions, memos were written with initial reactions to the participants’ responses. After the completion of a set of interviews, these were transcribed to text and collected with any other data. This data was analysed in detail, and memos were written on each section. The data was then coded, using the SuperCollider text editor, where codes could quickly be left as programmatic comments⁵. Once coding was complete, a script was run to extract all the codes from the text and place them in a mindmap using Freemind; this mindmapping software was a valuable tool for organising and analysing this data. With the codes collected together, similar codes were rationalised into single codes, with reference back to the data. Following this, the codes were grouped together into categories. The next stage was axial coding; this involved referring back to the data for each category and finding subcategories. These subcategories were linked together, creating a model of how the data was interrelated. Finally, this model was used to generate theories. Theories were based on stronger, saturated categories in the data, where there was enough discussion across participants to create a balanced and clear view of a topic. Weaker, unsaturated categories from the earlier studies were followed up and added to where possible in interviews in the later studies.

⁵The SuperCollider editor provides a hotkey combination to insert comments quickly between `/*` and `*/` symbols. It can also highlight comments so they stand out within the text, making the coding easy to read.

4.4 Summary

At the start of this chapter, some fundamental paradoxes of designing for creative tools were laid out: the need to design for unexpected use, and the importance of limitations and imperfections. This in turn affects the requirements for evaluating creative tools, and the manner in which they should be evaluated. Following this, an evaluation of the Wiimote as a musical controller was presented, as a case study testing HCI methodology. This study emphasised the importance of evaluating user experience over usability, and showed the strengths of collecting qualitative data for this purpose. The Wiimote study also demonstrated some of the shortcomings of lab-based evaluations, and it was proposed that real-world field evaluations could compensate for these, revealing insights into natural use of a controller and reflecting genuine creativity rather than use in contrived scenarios. As a mechanism for evaluating qualitative data from these studies, grounded theory was proposed as a solution; this technique has shown promise in other studies of creative technology, and is well matched to the aims of building an emergent picture of a tool's use. Putting this together, an evaluation plan was put forward, where new musical controllers are evaluated in a laboratory setting, which paves the way for a further field study, all using grounded theory for data analysis. These techniques were used for the evaluation of the two new musical controllers presented in this thesis, and their efficacy will be reflected on in the conclusions (see section 8.2).

This marks the end of the background and motivation in this thesis, and signals the start of the practical work, the creation of interfaces that address the problems laid out in chapter ???. Two systems are presented, both of which focus on interpreting hand motion for the control of digital music tools. The first system is *Phalanger*, a computer vision based hand tracking system.

CHAPTER 5

PHALANGER: A HAND TRACKING INTERFACE

‘In the front seat of Minister Srinivas’s car, Shaheen Badoor Khan slips his ‘hoek behind his ear. Taxiways, planes, airbridges, baggage trains merge with the interface of his office system. ... A flick of the finger *yeses* that report on Bharat’s combat readiness problem, *nos* that press release on further water restrictions, *laters* that video conference request from N.K.Jivanjee. His hands move like the mudras of a graceful Kathak dancer. A curl of the finger; Shaheen Badoor Khan summons the notebook out of thin air. *Keep me advised of developments re: Sarkhand Roundabout*, he writes on the side of an Air Bengal airbus in virtual Hindi.’

Ian McDonald (MacDonald, 2004)

5.1 About Phalanger

The use of hand movement for musical control has strong potential for enhancing musicality in interaction. Hand movement is natural and direct, giving the potential for easy learnability in a well designed system. It can also convey subtle, complex and fluid movement, with a huge potential for expressiveness and virtuosity in interaction. The motivation behind this project was to explore how hand motion could be tracked and used for musical control. Phalanger¹ takes a computer vision and machine learning approach to hand tracking and mapping, enabling markerless hand tracking and continuous control of musical parameters from changing hand geometry. It is also trainable, allowing the user to teach it to respond to customised sets of hand positions, thereby creating their own language of control and their own customised instruments. This chapter describes the design challenges behind building this system, the choice of hardware and implementation of software, and finally two evaluation studies. These were a formative laboratory based study, which laid the ground for a further field study. To begin with, research related to this project will be explored.

5.1.1 RELATED WORK

Many musical computer vision based projects have focused on large scale bodily gestures, for example iMaestro (Ng and Nesi, 2008) and EyesWeb (Camurri, Coletta, Varni, and Ghisio, 2007). This project however focuses on smaller scale hand motion, aiming to track detailed and subtle movements. A range of systems have been devised for finger and hand tracking using a variety of sensors and accessories. Some of these systems use markers or wearables, for example Wang and Popović (2009) use a colour patterned glove. Phalanger belongs to the group of systems that track the hand on its own, with no accessories or wearables. These systems tend to employ a combination of computer vision analysis algorithms and machine learning techniques to extract information from a video source and translate it into control data. Zhou, Xie, and Fang's (2007) Visual Mouse employs the Scale-invariant Feature Transform algorithm along with Principal Component Analysis to detect and track fingertips. Oka, Sato, and Koike (2002) use an infrared camera to track fingertips and fingertip gestures using a heuristic algorithm along with Hidden Markov Models. Premaratne and Nguyen (2007) use moment invariants as input to a neural network to recognise hand positions for control of consumer electronics devices. A more advanced system has been built by Agarwal, Izadi, Chandraker, and Blake (2007), who use stereo cameras along with a Support Vector Machine (SVM) to detect fingertip location and to distinguish between touch and hover positions. In a musical context, Burns and Wanderley (2006) have used the Circular Hough Transform to track a guitarist's fingers over frets. Also, Oliver (2010) provides a good example of computer vision hand tracking for musical performance. Phalanger combines features from these systems by focusing on broader scale hand motion together with more detailed finger motion, along with hand pose recognition.

¹The name is an inexcusable pun, referring to the *phalanges*, the bones that form our fingers

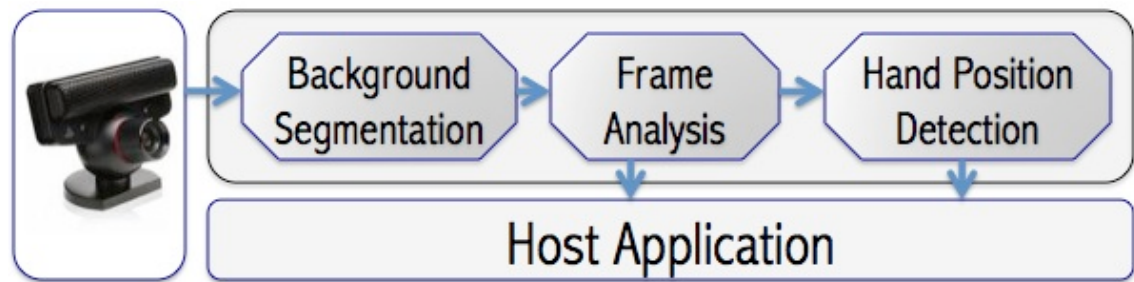


Figure 5.1: An Overview of the Phalanger System

5.1.2 IMPLEMENTATION

The main challenge of a hand tracking system such as this is in how to detect and track the subtleties of movement in a robust and reliable way. Along with addressing this challenge, Phalanger had some more practical design aims. One aim was to design a system which would work without wearables such as markers or sensors on the hand, which can be inconvenient and may impede motion. Another aim was to design a system which would work on lower cost and accessible hardware, thereby increasing the potential for use of the system and also increasing the possibilities for evaluating it. Phalanger was developed using the `openFrameworks`² C++ library, chosen for its range of add-on libraries, cross-platform portability and speed, which is essential for computer vision processing. It also uses the `openCV`³ computer vision library, along with `Fann`⁴ and `libSVM` (Chang and Lin, 2001) for machine learning. The reference system here is a MacBook Pro 2.2GHz laptop. For the video source, good results were achieved with both a Sony DCR-TRV80E firewire camcorder and a low cost (£25) Sony PS3Eye USB camera at 320x240 resolution. Processing occurs in three phases (see figure 5.1); firstly background segmentation, then frame analysis and finally hand position recognition. With regard to software architecture, the hand tracking code is built as a library which can be integrated within a host application.

5.1.2.1 BACKGROUND SEGMENTATION

The background segmentation process uses skin colour detection to separate the hand from the background. A neural network technique was chosen here so that the system would be dynamically configurable for each user's particular camera, room lighting and skin tone. Phalanger takes snapshots of the room without the user, and then of the front and back of the user's hand; these images are used as training examples for a back propagation network. Following from Mohamed, Weng, Jiang, and Ipson (2008), the pixel values are converted from RGB to the YCbCr colour space; in this way the luminance value (Y) can be discarded, leaving the chrominance values as neural network inputs. This makes the algorithm more robust to lighting changes, and allows for a smaller network. The network architecture was determined experimentally, and consists of two input neurons, four hidden neurons with linear transfer functions and one output neuron with a sigmoid transfer function. In use, the trained network

²<http://www.openframeworks.cc/>

³<http://sourceforge.net/projects/opencvlibrary/>

⁴<http://leenissen.dk/fann/>

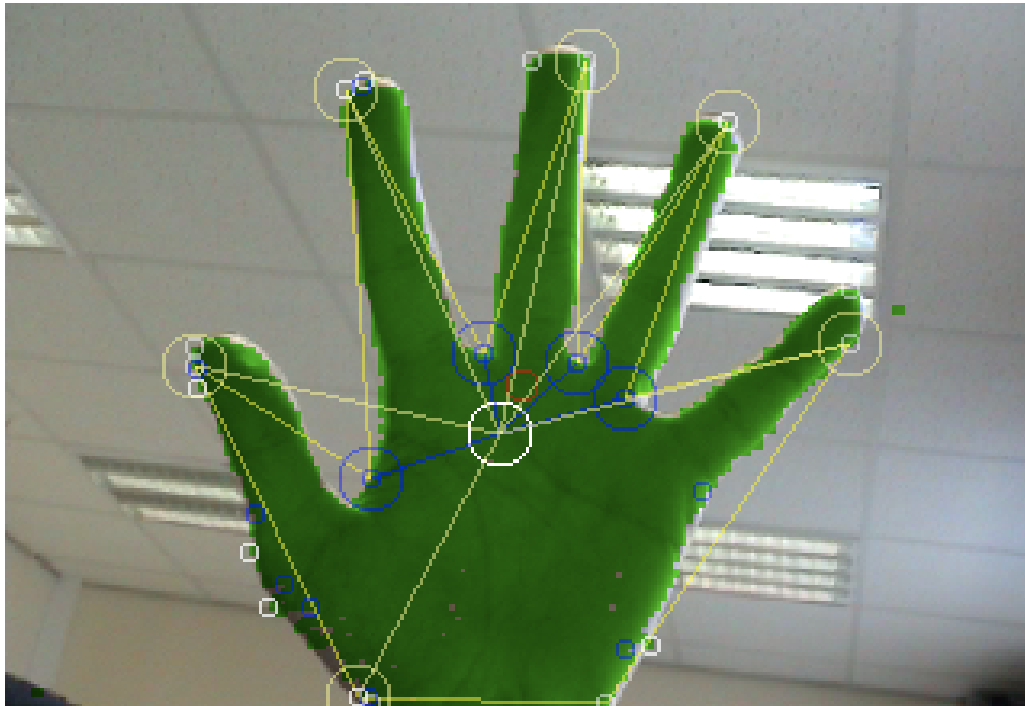


Figure 5.2: Frame Analysis in Phalanger

is run on every pixel of video data, separating the skin from the background as in figure 5.2.

5.1.2.2 FRAME ANALYSIS

The next stage analyses the video data with computer vision techniques implemented within the OpenCV library. There are an extensive selection of algorithms available for image feature analysis (Forsyth and Ponce, 2003); the following were chosen with the aim of arriving at a feature vector describing the hand which could be passed to a machine learning process. The machine learning algorithm that was used for this stage of processing required a fixed number of inputs, so it was also a requirement that this feature vector was of a fixed size. This was an interesting implementation challenge as the chosen computer vision algorithms output variable size data sets.

The process works as follows (the openCV functions used are shown in *italics*):

1. Grayscale output from the skin detector is smoothed out at the edges with erosion (*cvErode*) and dilation (*cvDilate*).
2. A blob and contour detection algorithm (*cvFindContours*) locates the rectangle enclosing the hand and detects the shape of the hand. The system makes the assumption that the largest blob on the screen is a hand and tracks this.
3. The hand contour is simplified to an approximation with a reduced number of data points (*cvApproxPoly*).
4. The simplified contour is used to find the convex hull of the hand (*cvConvexHull2*).
5. Convexity defects are derived from the convex hull (*cvConvexityDefects*).
6. The approximated contour and convexity defects are sorted in relation to their distance from the hand blob's centroid to obtain the six farthest points on the contour and the four nearest defects.

Table 5.1: Outputs of the Frame Analysis Process

6 furthest contour points from the centroid (normalised)
4 closest hull defects to the centroid (normalised)
angles between the far points
angles between the near points
first 6 Hu moments
the centroid (normalised)
height:width ratio of the hand blob rectangle

7. Finally, the contour is analysed to find the first 6 Hu moment invariants (*cvHuMoments*) (Hu, 1962).

Outputs from this process, summarised in table 5.1, are used both as inputs to the hand position recognition process and as tracking data for individual points on the hand. Figure 5.2 shows an example of an analysed frame. The white circle in the middle of the hand represents the centroid. The yellow circles on the finger tips and on the heel of the hand are points on the contour of the hand which are furthest away from the centroid. The blue circles are convexity defects taken from the convex hull.

5.1.2.3 HAND POSITION DETECTION

The hand position detection process is centered around a SVM, which observes the inputs from the frame analysis and attempts to predict the shape of the hand. Phalanger uses libSVM's *C_SVC* type SVM, with a radial basis function kernel configured with $C=2$ and $\gamma=0.2$. SVM workflow is similar to that of neural networks in that there is a training phase preceding a simulation phase. To create the training data, the system records feature vectors, over a number of frames, of the hand in one or more positions which represent a particular class. The number of frames needed for training varies with the overall number of classes and the quality of the training data. In some cases Phalanger works reliably with 10 frames per class, but generally recording 100 or over yields better results. Under-training the system can result in undesirable jitter in the output of the SVM.

The training process is choreographed and performed by the user of the system; it's human driven and therefore open to variability and error; the quality of training data correlates with the skill of the user at training the system. To aid the user during training, it was possible to display the realtime tracking information inferred from the video feed, and the system also showed the number of training images collected for each class of hand position. To obtain optimum reliability, a set of guidelines was developed to help create high quality training data, and these were passed on to users of the system in a later field study. The guidelines were as follows:

1. When recording data for a class, try and move your hand into all the possible positions you think will need to be in this class.

2. Make sure you don't overlap data by having the same poses in two classes. This will confuse the training process.
3. Move your hand around the camera field; there may be slightly different lighting patches or camera perspectives which you need to add into the training set.
4. Do not move your hand off camera, this will result in bad training data.

5.1.2.4 IN USE

The system is embedded within a host application, which can use the combination of hand position class and streams of hand geometry data as control data, either for direct musical control or through its own abstractions. Conditional logic based on the hand pose class can allow the host to infer extra information from geometry data. For example, if it is known that the hand is in a closed fist position with the index finger extended, it can be inferred that index finger tip position is the highest point in the contour. The applications implemented for the studies below provide examples of range of usage scenarios. This prototype version of the system achieved a speed of between 14 and 18 frames per second, and could differentiate between 8 classes of hand pose.

5.1.3 EVALUATION

The system was evaluated over the course of two separate studies. Initially a lab-based interview study was run (in late 2008) to obtain initial feedback about the system. After this, improvements were made to the system and a field study was undertaken (in early 2010), allowing musicians to use the system in their own environment for their own creative projects. The combined results of these studies resulted in valuable insights into both *Phalanger* itself, and into wider issues surrounding the use of this type of interface for digital music.

5.2 Formative Evaluation

At an early stage in the project, an informal 'formative' (Preece et al., 2002) evaluation seemed most appropriate, with the aim of acquiring initial feedback on which to base the next stages of development. Ten musicians (three female, seven male, aged between 18 and 30) took part in the study. They were all university students, nine of whom were studying Music Informatics. All had experience of using digital music controllers and software. They were asked to try out *Phalanger* in three different scenarios, giving their feedback in semi-structured interviews which broadly focused on their experience of trying out this style of interaction. Participants tried the system in three usage scenarios:

1. A Tetris like sound game (pictured in figure 5.3), where participants could knock falling blocks with their index finger to trigger different sounds depending on where the blocks landed. The blocks could also be stopped from falling by placing a finger horizontally underneath them.
2. Controlling sound with hand shape; the position of and angles between points on a path drawn around the extremities of the hand directly controlling granular synthesis parameters (pictured in figure 5.4).

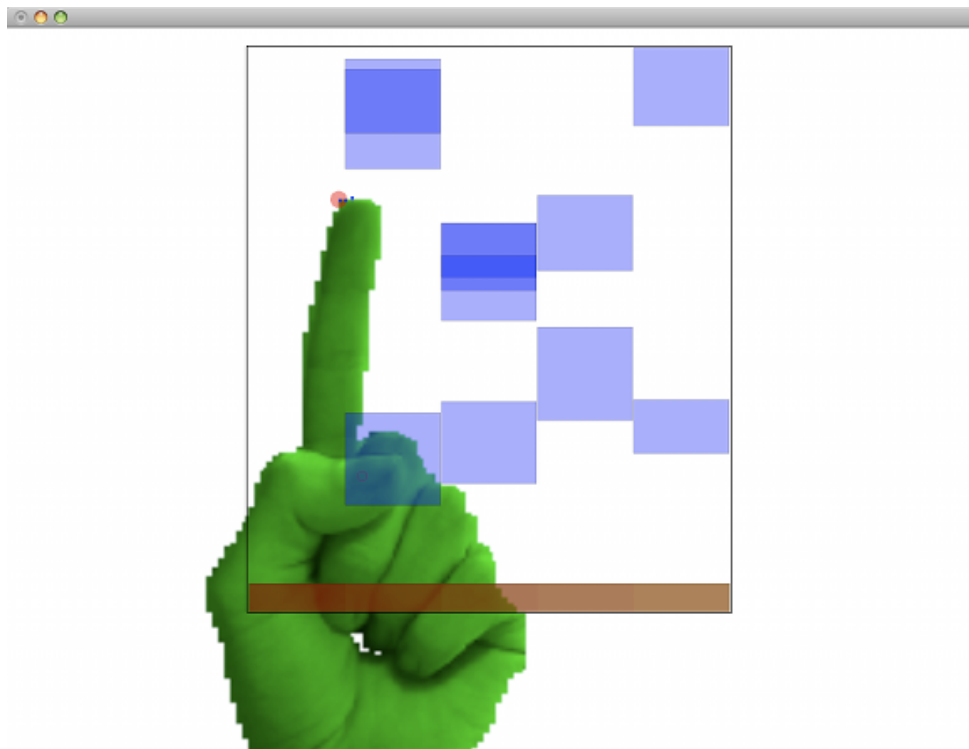


Figure 5.3: Evaluation Scenario 1: A Tetris Style Sound Game

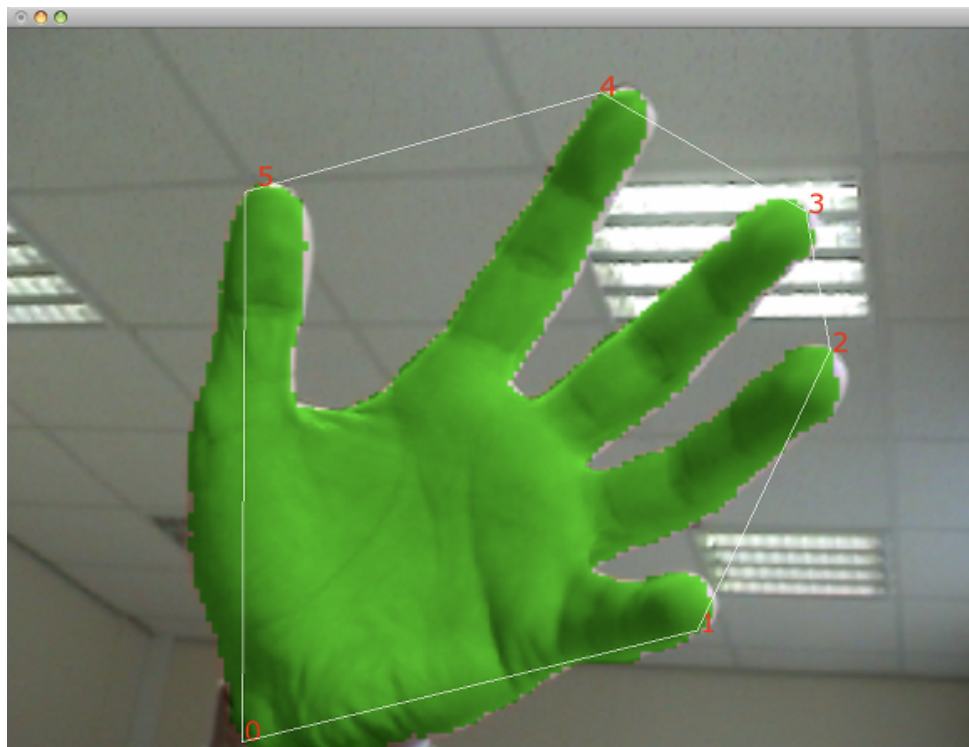


Figure 5.4: Evaluation Scenario 2: Hand Shape Mapped to Granular Synthesis

3. A sound mixing scenario (see figure 5.5) where the participants could navigate across a set of virtual sliders by moving their hand in a parallel plane to the camera, and zoom in and out by moving their hand forwards and backwards. By changing to a grabbing position, they could change the level of the sliders.

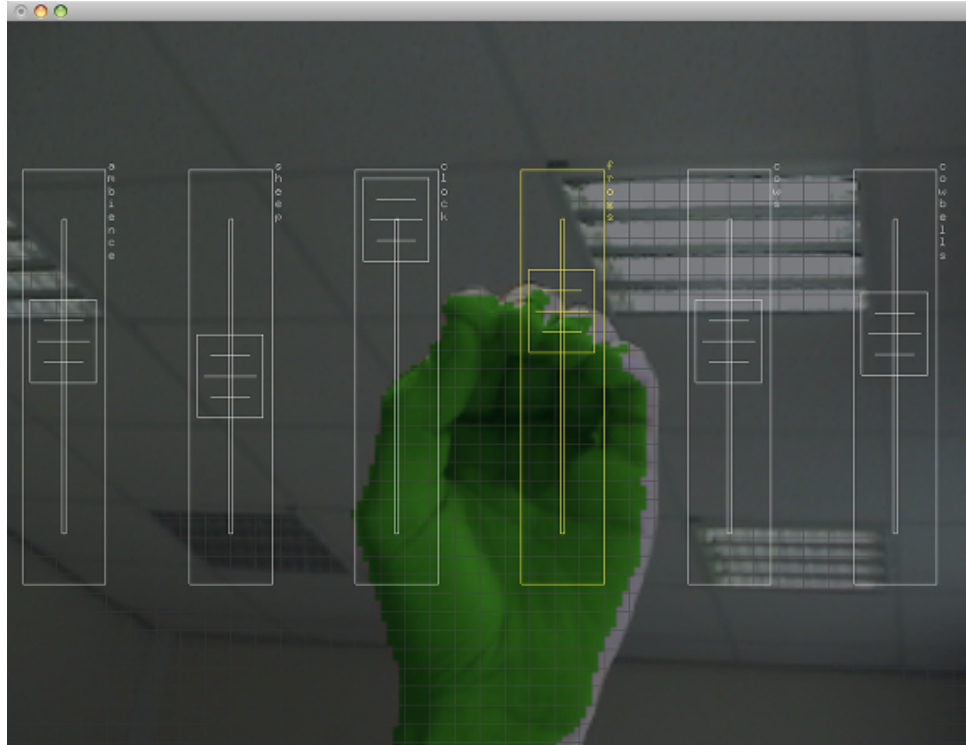


Figure 5.5: Evaluation Scenario 3: A Virtual Mixer

In all these scenarios, users controlled the software with their hand in mid-air facing a camera pointing up from the table, their elbows resting on the table. The scenarios were designed to explore the range of ways in which the system might be used in different musical contexts, testing mappings for discrete and continuous control, and both hand and fingertip motion. The interviews were analysed using a grounded theory approach (see section 4.3.5.2).

5.2.1 RESULTS

Responses fell broadly into four categories: control, feedback, ergonomics and learnability. In terms of control, reactions ranged from negative⁵ to unsure⁶ to positive.⁷

Half of the interviewees felt that the system needed to be more responsive⁸. Precision was also an issue for some,⁹ leading to suggestions for creative uses which suited less precise control¹⁰. There was some comment about the mapping of hand motion to sound; one interviewee described how they would prefer discrete gestures to continuous control for navi-

⁵I keep on moving things when I don't intend to', 'it's a bit unpredictable'

⁶I felt in control to a certain extent, I don't think I could quite find the direct correlation'

⁷it's got a really light kind of feel to it, I think I find it controllable', 'it's easy to control'

⁸it needs to be a bit faster somehow', 'speed of response could be better'

⁹fine control is difficult'

¹⁰I'd like to draw parametric EQ lines with hand movements whereas something like volume levels needs more precision'

gating the mixer, another participant liked the way that hand motion in the granular synthesis scenario seemed to match the way the sound changed.

As with the issue of controllability, there was a wide range of reactions about learnability. Some found the system difficult to adjust to¹¹ while others picked it up successfully.¹² This was related to issues of familiarity with a style of interaction most had not tried before.¹³

There were some interesting results concerning visual feedback which again showed the range of preference between the participants. *Phalanger* shows the user's hand on the screen in three variations: the hand on its own, the hand with markers from the frame analysis, and the markers on their own. All three modes were preferred by different participants, and in some cases their choice of visual feedback made a significant difference to their ability to control the software. In the granular synthesis scenario, one participant felt comfortable with the screen turned blank.

Some of the participants discussed physical feedback,¹⁴ emphasising the absence of physical resistance in the interface.

Finally, ergonomics was a strong theme. Some hand movements didn't seem natural,¹⁵ and some people felt fatigued, their fingers, arms or shoulders tiring during use of the system.

5.2.2 DISCUSSION

The evaluation results highlighted some areas where improvements needed to be made to the system. Responsiveness was a key issue, signalling that optimisations needed to be made to improve the frame rate of the system. The ergonomics results show that the hand position used in the evaluation could be tiring. A better solution would prove to be using an overhead camera with the hand resting on a horizontal surface. The results also show how some hand and finger movements can seem unnatural, something to be kept in mind when specifying control movements for this system. Of particular interest was the range of user reactions, some feeling instantly comfortable using the system and some finding it harder to use.

5.2.3 SUMMARY

The initial version of *Phalanger* reached its initial objectives of providing a system through which users can control music software with hand motion. The evaluation explored use of the system in a range of mappings and scenarios; the results highlighted some areas which need attention, gave insights into new design features which would improve operation, and helped to build a picture of how musicians respond to this style of interaction. The next phase of development was to implement these improvements and conduct a more detailed field evaluation of the system.

¹¹'sometimes I get it and then sometimes I seem to lose what I'm doing'

¹²'it's taken me a little while to get used to it, I'm finding I can quite consistently get it to do what I want now'

¹³'the mouse is easier because I've used it before', 'this is weird, I'm used to a mixing desk'

¹⁴'there's nothing to resist your movement, I don't know when I should stop my hand', 'there's nothing in terms of feedback, I'm not pushing anything'

¹⁵'waving your finger like this isn't the most natural thing', 'I find it harder to go that way left to right'

5.3 A Field Evaluation of Phalanger

While the formative evaluation elicited some valuable insights into the prototype controller, the artificial setting and timescale of the study did not reflect real-world use. To obtain a more detailed and realistic assessment of the controller in use, a qualitative field study was carried out. This took place in the participants' own creative setting, using the controller with their preferred software and materials, in their own time.

5.3.1 IMPLEMENTATION CHANGES

Following from the initial evaluation study results, a number of improvements were made to the system, to improve performance and user experience.

5.3.1.1 ERGONOMIC IMPROVEMENTS

Participants in the previous study complained of fatigue when using the system, set up as it was with the camera on the desk pointing up diagonally at the hand. The system was changed to work with the camera facing down on the desk, the user being able to rest their hand on the surface or suspend it in mid-air if required. To achieve this, a lightweight Sony PS3Eye camera was placed in a retort stand, giving flexible and accurate position adjustment. Figure 5.6 shows an example of the system in use like this.

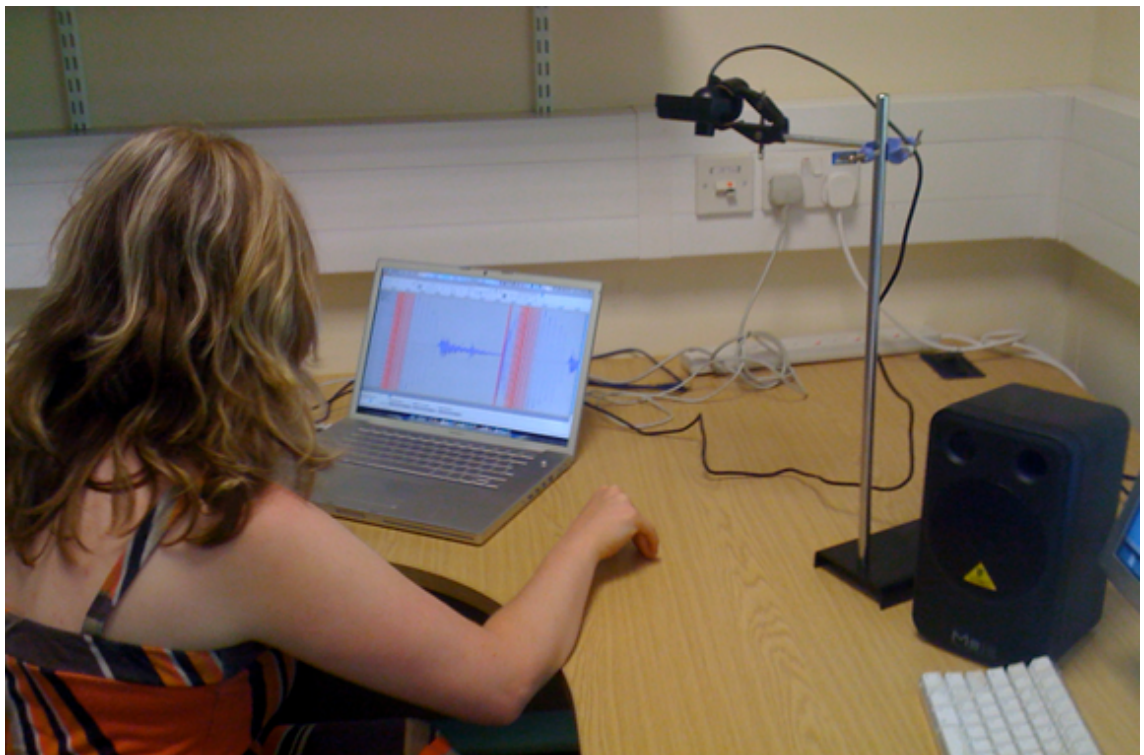


Figure 5.6: Phalanger In Use With A Desk Facing Camera

5.3.1.2 PERFORMANCE IMPROVEMENTS

The image processing was modified in several areas in order to improve speed and reliability. The skin detection algorithm was proving to be a bottleneck, so instead of running the colour detection algorithm on every single pixel, this was changed to detect the colour of one pixel in a 3x3 block, and assign this whole block the result from the single pixel. This improved performance at the expense of some resolution.

The other main change was in the hand pose detection. In the previous system, a feature vector was created from a selection of geometrical descriptors of the hand. In the new system, an image of the hand is taken from the blob surrounding it, and scaled into a 21x21 image, adding extra white space if necessary to preserve proportionality. This image is read as a 441 point feature vector, each datapoint being the greyscale value of the corresponding pixel normalised to between 0.0 and 1.0. Using a graphical feature vector instead of a geometric one made a significant improvement in the feel of using the system; jitter in the output of the SVM was reduced, therefore smoothing on the output could be reduced or removed so that the system then reacted much more quickly and accurately to changes in hand pose. With the performance improvements in place, the frame rate stays around 28-30fps in typical usage, dropping towards 25fps if the hand is moved closer to the camera.

5.3.2 FIELD STUDY PARTICIPANTS

Three participants took part in the study (one female, two male, aged between 20 and 40). All were musicians who were actively involved in the composition and performance of electronic music; one was a student of electronic music, one a teacher and researcher in computer music and the other a professional composer and performer. All participants were volunteers. Due to the busy schedule of participants, the study took place over varying periods, from six weeks to three months.

5.3.3 SOFTWARE

The software for the study is designed to integrate with existing music applications, working as a controller for a variety of different possible sound engines. One participant used SuperCollider (McCartney, 2002) as their main software tool, so a version of the Phalanger software was created that output data as Open Sound Control (OSC) (Wright, 2005) messages, allowing it to communicate with SuperCollider. The other two participants used Ableton Live¹⁶ as their main tool, so a variation of the software was created that sent out tracking data as MIDI messages, which could be mapped to continuous parameters in Live. These versions shall now be discussed in more detail.

5.3.3.1 COMMON FEATURES

Both the MIDI and OSC versions vary in the way they send out data to other applications; however for both versions the calibration and training processes are identical. To calibrate the skin recognition system, initially the user is guided through the sequence of collecting images

¹⁶<http://ableton.com>

of the background and their hand by a series of onscreen messages. Once calibrated, the user can control the software from the menu system or through key commands. Data collection for training hand pose recognition is collected by selecting the target class and then recording hand movements in blocks of one hundred frames. Once data is collected, the SVN training process can be initiated, after which the software will track the hand and send out data to accompanying software. The menu system also facilitates access to utility functions such as saving, loading, and training set management.

5.3.3.2 THE OSC VERSION

The OSC version communicates with any OSC compatible software, such as SuperCollider, Max/MSP, PureData and so on. It sends out the index of the current hand pose according to the training SVN, together with streams of geometric data describing the shape of the hand. The user can create algorithms and mappings in the accompanying software to process the control data in the manner of their choosing. Table 5.2 describes the OSC messages output by the software. All geometric values are normalised to between 0.0 and 1.0.

5.3.3.3 THE MIDI VERSION

Unlike most OSC software, a MIDI client such as Ableton Live cannot react conditionally to data changes such as the hand position index, so a system was added to the software to conditionally send the tracking data as MIDI controller messages, dependent on the current hand pose. MIDI clients are also usually incapable of performing mathematical processing on the data streams so two derived parameters were output together with raw data.

The system work as follows: for each individual hand pose, the system can optionally send out any of 12 different tracking variables on individual MIDI controllers; for the first hand pose, these variables (listed in table 5.3) are sent out on controllers 0-11, for the second they are sent out on 12-23, and so on. Using a GUI (see figure 5.7), the user can configure which streams are enabled, and the range in which they are sent. The control data is sent out over an internal IAC bus, and can be mapped as required in a MIDI client. For example, if the x coordinate of the hand centroid is enabled for hand poses zero and one, then the software will send out the centroid x on controller 6 if it detects hand pose zero, and controller 18 for hand pose one. In this way, a system of interaction can be built up where, for example, the hand modulates a reverb level with the hand in position zero, and then a delay send in position one. To prevent the values jumping between hand poses, an option was added so that the system could be limited to sending out MIDI data only when the space bar was pressed.

5.3.3.4 OTHER SOFTWARE

With both MIDI and OSC modes of interaction in place, Phalanger is capable of working in accompaniment with the majority of music software. Participants were also supplied with *Macam*¹⁷ (pictured in figure 5.8); they used this to fine-tune the camera settings for the environment within which they were using Phalanger, so that the skin recognition process could work optimally. Importantly, this software allowed the auto white-balance feature of the drivers to

¹⁷<http://webcam-osx.sourceforge.net/>

Table 5.2: Phalanger OSC messages

OSC Message Name	Values	Description
<i>/phalanger/handPosition</i>	<i>i</i>	The index of the current hand position, as recognised by the trained SVM. This message is sent only when this value changes
<i>/phalanger/handPositionSmoothed</i>	<i>i</i>	A smoothed version of the previous value, to avoid possible jitter when on the border between hand positions.
<i>/phalanger/centroid</i>	<i>x,y</i>	The x and y values of the centroid of the hand
<i>/phalanger/topmost</i>	<i>x,y</i>	The x and y values of the topmost point of the hand
<i>/phalanger/leftmost</i>	<i>x,y</i>	The x and y values of the leftmost point of the hand
<i>/phalanger/rightmost</i>	<i>x,y</i>	The x and y values of the rightmost point of the hand
<i>/phalanger/centroidToTopAngle</i>	θ	The angle between the centroid and the topmost point of the hand
<i>/phalanger/area</i>	<i>a</i>	The surface area of the hand (useful for pseudo-3d motion tracking)
<i>/phalanger/contour</i>	$x_1, y_1, x_2, y_2 \dots x_n, y_n$	A list of co-ordinates describing points on a contour around the hand, sorted by angle from the centroid
<i>/phalanger/defects</i>	$x_1, y_1, x_2, y_2 \dots x_n, y_n$	A list of co-ordinates describing contour defect points around the hand, sorted by angle from the centroid

Table 5.3: Phalanger MIDI output

Variable	Description
<i>centroid x and y</i>	The x and y values of the centroid of the hand
<i>topmost x and y</i>	The x and y values of the topmost point of the hand
<i>leftmost x and y</i>	The x and y values of the leftmost point of the hand
<i>rightmost x and y</i>	The x and y values of the rightmost point of the hand
<i>Centroid to top angle</i>	The angle between the centroid and the topmost point of the hand
<i>Area</i>	The surface area of the hand (useful for pseudo-3d motion tracking)
<i>Left to top distance</i>	The distance between the leftmost point and topmost point; useful for tracking thumb to forefinger distance
<i>Width</i>	The distance from the leftmost to the rightmost point

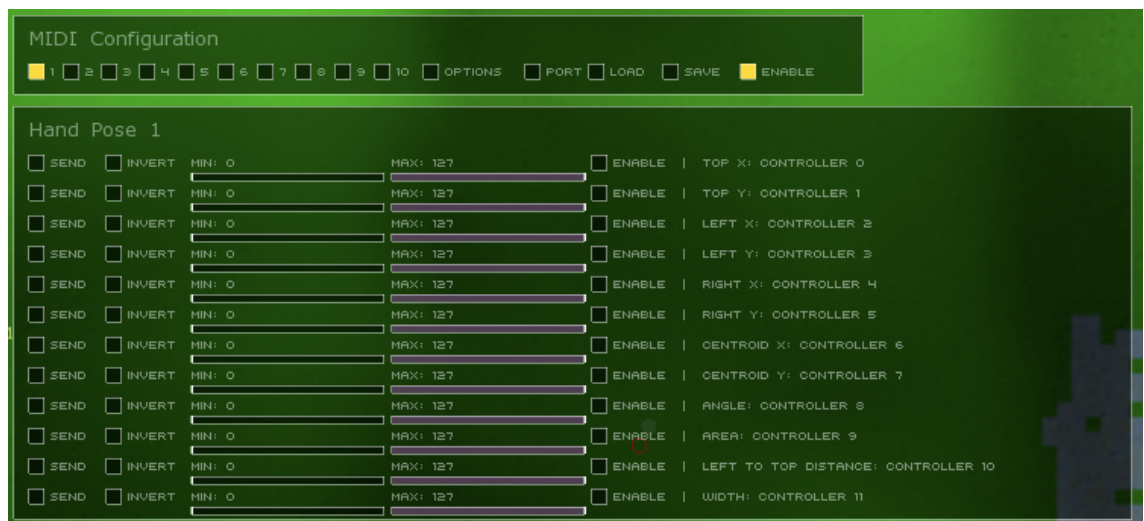


Figure 5.7: Phalanger MIDI Configuration

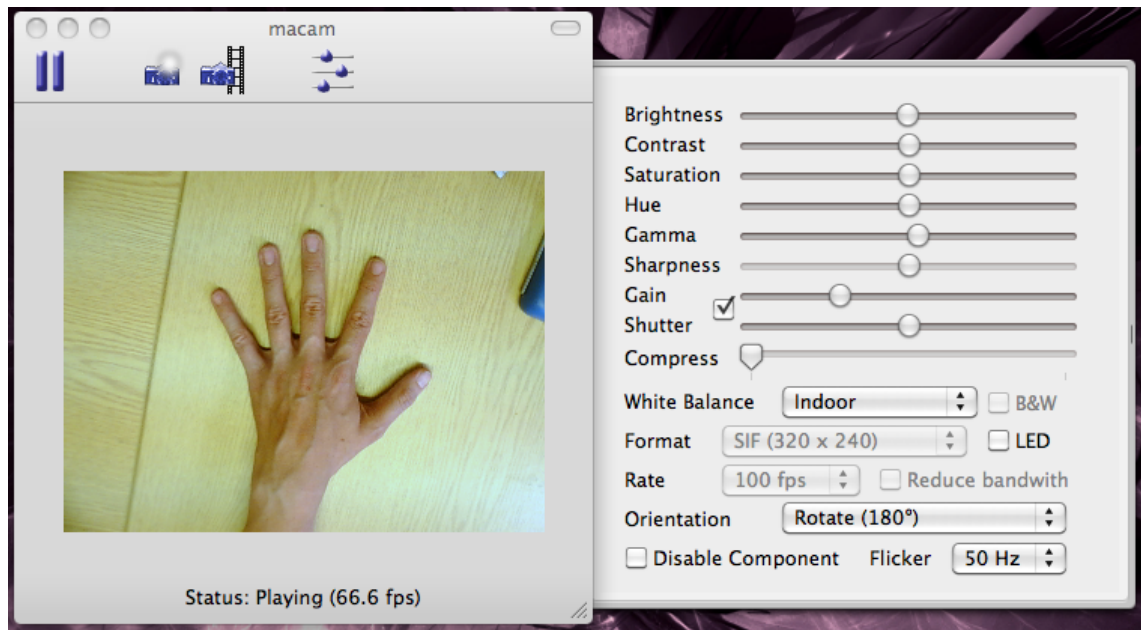


Figure 5.8: Macam

be disabled. Without this setting disabled, the camera would automatically adjust its gain and shutter settings with changing lighting conditions, possibly affecting the quality of skin recognition.

5.3.4 THE STUDY

The participants were loaned the hardware for the duration of the study, which consisted of a Sony PS3Eye camera and a retort stand which acted as a variable position camera mount. To commence the study, an initial face-to-face meeting took place where the participant was set up with the hardware and software, and given a tutorial on how to use the system. In one case, the software was customised to some specific requirements for the participant's desired configuration. The participants were supplied with a written manual for the system, and were told that they could get in contact at any time during the study with queries or problems. They were also asked to keep a record of their experience through the study, using any media they chose (such as a diary or video). After the period of evaluation, a meeting was arranged for a final interview. For two participants, this took place in person, for the other this was conducted over Skype. The interviews were semi-structured, and were recorded as audio.

The nature of use of the system during the evaluation varied considerably between participants. The participant who used the OSC version with SuperCollider was a computer music researcher, performer and instrument designer. They took an explorative approach, focusing on probing the constraints and affordances of the system, and looking at compositional aspects of use. The users of the MIDI version were both DJs, composers and live performers of electronic music. One of these users focused on performance aspects, controlling sound in Ableton Live. The other participant lacked time in their schedule and made limited use of the system, although they were still able to offer some valuable insights into its use.

5.3.5 DATA ANALYSIS

The following data was collected for analysis:

- Participant 1.
 1. A transcription of the final interview.
 2. A section of an article submitted to a journal, concerning a comparison between musical interfaces, one of which was Phalanger.
- Participant 2.
 1. A transcription of the final interview.
 2. A video of the participant performing a piece, during which Phalanger was used as a controller. A text was created listing detailed observations about the video.
- Participant 3.
 1. A transcription of the final interview.

The data was analysed with grounded theory techniques (described in section 4.3.5.2). The main concepts and categories discovered in the data are shown in figure 5.9.

5.3.6 RESULTS

Figure 5.9 shows a summary of the categories and concepts that emerged from the analysis. These are now discussed in detail.

5.3.6.1 CONCERNING PHALANGER

The participants had some common difficulties with using the system, both functionally and conceptually. The principle issue was with the calibration of the skin detector; Phalanger needs a reasonably constant lighting environment, both in terms of light being evenly spread across the camera area and in terms of the light not changing too much over time. There were initial problems with establishing the right environment for the system to work effectively, and with operating the training process for the skin detector. However, after becoming more accustomed to the system this became less of a problem. Participant one felt that these problems lessened their confidence in using the system, feeling that sometimes the calibration didn't work for them. This area of calibration is tied in with creative workflow; it limited the portability of the system, making it more difficult to use spontaneously. Another potential area of difficulty was with data flow; the system constantly streams multiparametric data, and participant three found this amount of data '*overwhelming*', being accustomed to the more conventional event-response paradigm of interaction.

The participants' view of the function of the system varied; participant three was interested in it primarily as a compositional tool, while the others viewed it as a performance system and hadn't considered it for use in a compositional context.

Designing and training configurations for the system was a central issue, and this shall be discussed in more detail later. Participants generally felt competent at the process of training

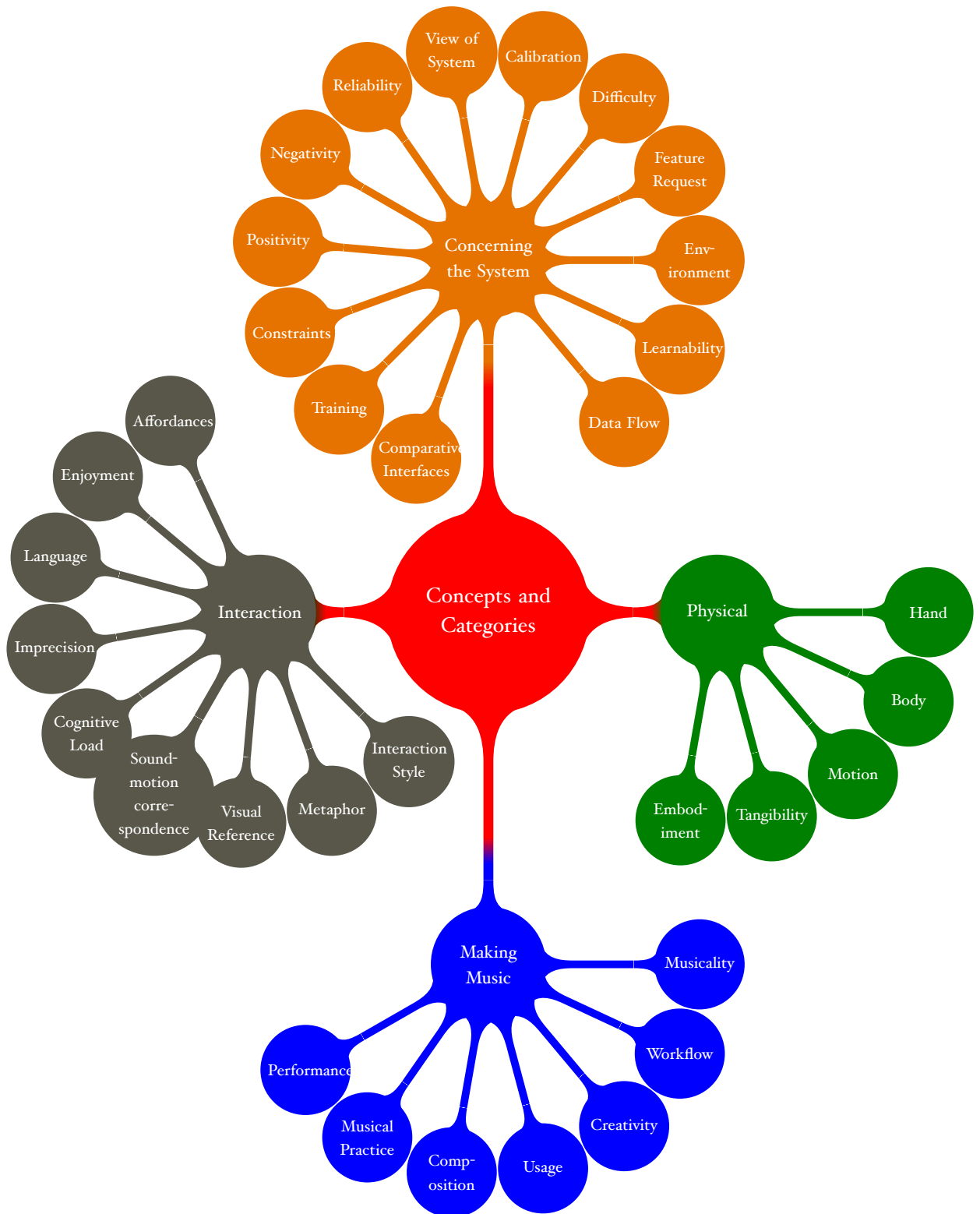


Figure 5.9: Phalanger Field Evaluation Interview Analysis: Concepts and Categories

the system to recognise hand poses, and this was a process that became easier over time. Overall, there seemed to be a manageable learning curve to the system. Performance with new settings required practice, as it would with any new instrument. One participant discussed how it was no problem to return to a configuration a few days later and be able to pick it up again.

All participants gave positive feedback about the system; it was enjoyable to use the hand expressively with the controller, and the novelty of the system was appreciated. Comments included *'I found the experience of using Phalanger very interesting, inspiring and rewarding. It is a well engineered and stable system.'*, *'I love the hand control element of it.'*, and *'You got to put a lot of expression into controlling the parameters, so it's really cool'*.

Potential new features for the system were discussed. This mainly centered around forms of discretised interaction, using gestures to trigger events such as notes. One request (from participant one) was for a meta-level mapping system, which would allow fast switching of mapping configurations, enabling different settings to be used over the course of a performance.

5.3.6.2 THE PHYSICAL

The physicality of using the system was a key issue. Unlike a conventional instrument, which provides tactile, material feedback, the Phalanger user makes use of proprioceptive senses to operate the controller, something not typically associated with computer use. According to one participant, this gave the controller an ethereal, disembodied feel; this invites comparison with theremin like instruments. The affordances are abstract, and more difficult to explore; you can't explore by touch so you need to think more about how the system works. With this lack of tangibility, one participant imagined tangible metaphors when using the controller, imagining that they were pushing particles around, or using a DJ deck when using a flat open hand posture. They said they felt a close bodily connection with the music, sub-consciously tensing or relaxing their arm and hand according to the music and the parameters they were controlling. The way in which the system allowed this kind of expressiveness in control was another positive point of use.

Some discussion focused on the hand. Focusing first on the limitations of the hand, there were comments that it felt unnatural to interact only with the hand rather than the whole body, and that it would have been preferential to be able to use both hands. The biological limitations of the hand were acknowledged, the amount of different poses the hand is capable of defining the limitations of how the system can be configured. Connected with this is the user's individual capacity for memorising hand poses, and creativity in selecting hand poses for new mappings and configurations. Participant two commented on how, during performance, they never felt confused between hand poses, and felt comfortable switching between them. In general, it was felt that the system was intuitive to use, participants commenting that it was instinctive and natural.

5.3.6.3 INTERACTION

The key themes in the area of interaction were language, cognitive load, and imprecision. When training the system with hand poses and designing mappings for these hand poses, you are creating a language of interaction, which is different for each new configuration. This highlights a number of issues. Firstly, creation of this language proved problematic for participant three, who felt they had difficulty imagining different hand poses to use with the system. They also felt that hand poses were not necessarily symbolic with the music; interaction felt more abstract than, for example, using a knob. This extra level of abstraction, together with the need for remembering the different shapes, added to the cognitive load of using the system. For another participant, no such problems existed; they felt that choosing hand poses felt natural and that they didn't have to think about changing hand poses when using the system.

Imprecision was something that all participants commented on; being operated with proprioceptive senses, the participants felt there was potential for error when using the system. This is a trade off with having many degrees of freedom. This potential for error, or perceived lack of security, affected participant's confidence in using the system, especially as it was a departure from their normal musical practice of using precise hardware control. However, the freedom of movement was considered a very positive attribute of the system, participant three comparing this freedom to the freedom for error in playing an acoustic instrument.

Participant two, who made a video of their performance with the system, discussed how they rarely looked at the hand while using the system, relying on audio feedback to detect if their hand had left the boundary of the camera field. This is confirmed in their video where they look at the laptop screen while using *Phalanger*, only occasionally looking over to the hand.

5.3.6.4 MAKING MUSIC

Participants discussed when and how they made music with *Phalanger*, together with the creativity and musicality aspects of using the system. The system presented some barriers to creative workflow; first, there is a preparation time before the system can be used where the skin detection must be calibrated. There is also a potential for mid-session interruption if the lighting conditions change significantly. These factors, coupled with portability issues mentioned earlier, potentially affected the frequency of use.

There were two sides to the creative use of the system; creativity in the design of interaction and mappings, and then creativity in performance. The flexibility of configuration and usage allowed the emergence of a personal style of performance. Overall, one participant described the system inspiring in use, especially given its constraints. Another felt that it enhanced their creativity.

During the study, participants mainly created mappings for the continuous control of synthesis parameters, although participant three also discussed how they used discrete mappings, using a change in hand pose to trigger an event. Participant three, emphasising the transient nature of their usage of the system, discussed how they used different mappings in each session, each session prompting new ideas. Participant two, using the system for a performance, returned to the same configuration in consecutive sessions.

In terms of musicality, comparisons with the nature of musical instruments were made. Continuing the theme of transience, participant three commented that the system felt like a new instrument with each new setting; in effect training and mapping is akin to designing a new instrument. While participant two felt that the controller shared qualities with an acoustic instrument in terms of its expressive potential and potential for imprecision, participant three felt that it was more akin to an '*HCI interface*' than a musical instrument.

Participant two felt that hardware controllers had limited expressive potential and that this kind of interaction allowed more expressive freedom. In their video, musically expressive motion can be observed; their hand and arm tensing with the music, and the use of expressive variations of hand poses.

5.3.7 DISCUSSION

5.3.7.1 SCOPE AND LIMITATIONS OF THE STUDY

While the study has resulted in some rich data, the scope and limitations of the study and results must be acknowledged, in terms of the time the participants spent with the system, and the depth of use. Two out of the three participants felt that they would have liked to spend more time with the system, but couldn't for various reasons. One participant in particular had trouble fitting the study into their schedule, although their insights into the system were still valuable not only because they did spend some time exploring it, but also because they provided the expert opinion of an experienced electronic musician. None of the participants made full use of all the features available, so the results here illustrate exploration of the basics of the controller rather than in depth use. Despite these shortcomings, the results provide a compelling picture of the controller *in the wild*, being used creatively by musicians and providing data well beyond the scope of a lab-based study. Bearing these factors in mind, a discussion of the results follows.

5.3.7.2 DISCUSSION OF RESULTS

The results paint a picture of a relatively complex interface that works on three levels, requiring some level of skill in each. Firstly, there is the functional calibration stage, at which the user needs a level of competence to establish a basic level of functionality. Following this there is training and then use, each involving different kinds of creativity, firstly in defining a language of interaction, and then in using this language to interact musically. The creation of a successful language is fundamental in a system such as this, and the key issue here is establishing a meaningful connection between motion and audio output. The importance of this relationship is explored by Antle et al. (2009), who found that use of an embodied metaphor could make a system more intuitive to use. This is echoed in the evaluation results; one participant had difficulty making a symbolic connection between motion and sound when using the system, and commented on how the extra cognitive load of using a new vocabulary of gestures made the system more difficult to use. Another participant, who imagined metaphors when interacting with the system, found the system intuitive and natural. In Phalanger, this connection between motion and sound is established over a two level mapping process, first by defining gestures, and then by defining mappings between the system output and the audio

engine. The conceptual split between these two stages may contribute to some difficulties in use of the system.

Returning to the wider research theme, the results highlighted some issues surround the musicality of the interface. One participant enjoyed the potential for error in the interface, feeling it contributed to an instrument-like feel. This theme of imprecision and unpredictability is also observed in other studies of computer musicians (Bertelsen et al., 2007; Magnusson and Mendieta, 2007; Gelineck and Serafin, 2009b), and this highlights the importance of incorporating these factors into musical interaction design. Participants commented on the expressive potential of hand use; a system such as this has an inherent convergent mapping both on a physical and logical level, translating from the many degrees of freedom of hand motion into a lower dimensional set of continuous parameters. Wanderley and Depalle (2004) suggested that convergent mappings have a high potential for musical expressivity, and this is confirmed in the results. The interplay between this high degree of freedom and potential of error underpins the intangible style of interaction in *Phalanger*.

5.4 Summary

A system has been presented that enables musical control with hand motion, achieved by tracking the hand with a combination of computer vision and machine learning techniques. A formative, laboratory based evaluation revealed potential improvements to the system; these were implemented and the system evaluated in more depth in a qualitative field study. The results showed that hand motion tracked in this manner can be successful in creating expressive, musical interaction when there is a meaningful relationship between hand motion and sound. The creation of this relationship presented some potential difficulties, and questions remain open in the area of how to design a language of interaction, and how this relates to creative workflow, cognitive style (Eaglestone et al., 2008) and the musical practice of the user.

The *Phalanger* system was explored in a further study, as an interface for a timbre space exploration system. This work is presented in chapter 7. The study elicited more data about user experience with this system. Along with the formative and field study results, this gives a total of three sets of results from which to triangulate an understanding of *Phalanger*. These results shall be explored together in chapter 8.

CHAPTER 6

ECHOFOAM: A MALLEABLE CONTROLLER

‘It’s got to the point now where my work
is really about riding that knife-edge
between what works and doesn’t work,
absolute control and no control.
Disaster and delight. Life, really!’

Kaffe Matthews (Rodgers, 2010)

6.1 Introduction

The EchoFoam system is a new controller that further explores the nature of embodied control of digital music tools, following a multiparametric paradigm. While Phalanger explored freeform hand motion and intangible interaction, EchoFoam looks at different aspects of motion capture, using a tangible interface. The system follows a paradigm of manipulating malleable material in order to afford detailed and intuitive control.

Malleable interfaces are currently seeing commercial outings in the form of the SUMA¹ and Blobo² devices, both marketed as *squeezable* controllers. There has also been plenty of activity in the academic sphere. Schwesig (2008) describe the concept of *Organic Interfaces*, sensitive analogue devices that acknowledge the subtleties of physical interaction; they illustrate their ideas with the hypothetical Gummi device, a deformable display that responds to physical manipulation by the user, arguing that subtle physical interaction with a real-world object such as this would lead to a suspension of disbelief, a quality perhaps also desirable in a musical controller. Moving to real-world examples, Reed (2009) created a prototype digital clay, using embedded wireless sensors and computer vision to measure manipulation of the material. Sato, Mamiya, Koike, and Fukuchi (2009) also use computer vision techniques in their *PhotoelasticTouch* system, which measures the deformation in transparent rubber objects in a tabletop interface. An example very relevant to this project is Smith et. al.'s work; they created several input devices using configurations of multiple conductive foam sensors, for use as interfaces for 3D sculpting and camera control (Smith, Thomas, and Piekarski, 2008b,a). Milczynski, Hermann, Bovermann, and Ritter (2006) created the *Elastable*, a device that employs computer vision to measure deformation of a rubber surface in order to explore and sonify high-dimensional data sets.

Focusing on musical examples, a novel route to malleable control is taken by Hook, Taylor, Butler, Villar, and Izadi (2009). They used an array of ferromagnetic sensors to measure deformation in a ferrofluid filled bladder, in one example mapping this to synthesis parameters in Max/MSP. Chang and Ishii (2007) designed the ZStretch musical controller, a fabric device that measures deformation using resistive strain transducers sewn into lycra. Marier (2010) embedded accelerometers and force sensing resistors into a piece of sponge to create a malleable musical controller. Lastly, another musical example is Weinberg and Gan's (2001) *Squeezables*. They embedded pressure sensors into several soft gel balls which were played together as a collaborative instrument. An evaluation study showed that the players valued the expressiveness of this style of interaction.

The system described here uses a malleable foam sensor, tightly coupled with reservoir computing mapping techniques, to create a device that can measure subtle physical manipulations and map them to multi-dimensional control streams. It is used in this case for the haptic exploration (Lederman and Klatzky, 1987) of sound synthesis parameter spaces. Two user studies, one lab-based and one field-based, evaluate the success of this system and highlight some interesting issues surrounding malleable control.

¹http://www.cambridgeconsultants.com/news_pr257.html

²<http://www.bloboshop.com/>

6.2 EchoFoam

Considering malleable control in the context of the wider design framework, the ability to track subtle, detailed physical manipulations was a key objective, and the choice of sensors was the most important decision in achieving this. Inspired by Smith et al.'s (2008b) controllers, conductive foam was chosen as the sensor material. It is also known as Electrostatic Discharge (ESD) foam, and is commonly available as a packaging material for ESD sensitive electronics components. This foam can be appropriated as a material to create pressure sensors. In this project there is a key difference to Smith et. al's projects; where they used multiple independent foam sensors, this project uses a single piece of foam as a continuous sensor, with multiple contacts measuring the state of deformation. In this way, both large and very subtle changes in the shape of the sensor can be detected with a relatively simple electronic circuit design, giving the sensor high potential for expressivity in control. The measurements from this kind of sensor are complex and interdependent, so the system is tightly coupled with echo state network mapping techniques to create usable control streams from the output voltages. The system was named *EchoFoam*.

6.2.1 SENSOR CONSTRUCTION

Conductive foam has a key property for making a malleable sensor; its electrical resistance changes when deformed. To exploit these properties, a current can be applied to a piece of foam and the resistance measured between multiple points within it. When the foam is deformed, the resistance changes in the area of deformation and the readings change accordingly, giving a consistent and localised measurement of physical manipulation. If the resistance is monitored at a wide range of locations in the foam, each individual deformation of the foam as a whole will be identifiable by a consistent set of readings, and even very small deformations will be detected.

The sensor is constructed using low density conductive foam. 32 swg enamelled copper wire, selected for its thinness and flexibility, is tied into the foam with the enamel scraped away at the ends to make electrical contacts. One wire is used as a live contact at 5 volts, while the rest of the wires measure resistance, running to voltage dividers on a breadboard, and then into eight-channel analog multiplexer chips (type 74HC4051) controlled by an Arduino board (Banzi, 2009). A control program running on the Arduino scans the voltage from each wire with the multiplexers and sends them to a computer via USB at 115200 baud. In this configuration, 16 sensor wires were used (see figure 6.1). The wires were tied into 8 foam squares which were glued together into a cube shape. The system takes measurements from the sensor at approximately 100Hz.

A key design decision concerned how to place the wires in the foam in order to give the most useful data readings. Measuring resistance at multiple points in a single piece of foam leads to a stream of interdependent measurements. Due to the complexity of this information, the absolute position of the foam cannot be tracked through perceptually meaningful single values; instead the set of readings taken as a whole gives a consistent signature for any particular deformation. This means that the wires did not need to be placed in precise ge-

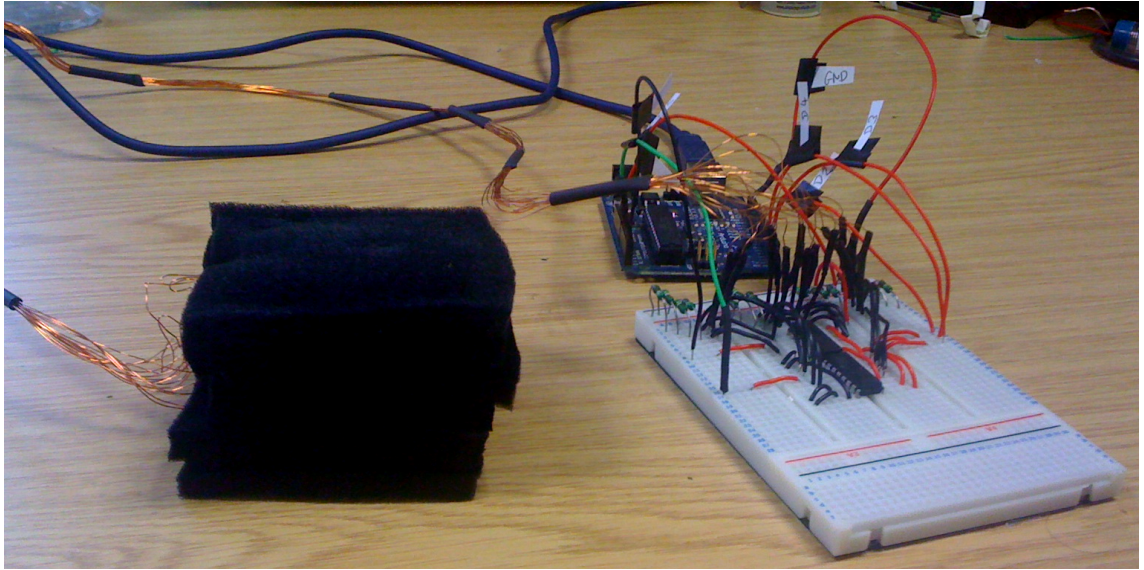


Figure 6.1: An EchoFoam Controller

ometric positions in the foam, although it was important that they covered all parts of the sensor. With this in mind, the wires were placed at irregular points in the foam, such that as wide an area of the foam as possible was covered.

6.2.2 MAPPING WITH ECHO STATE NETWORKS

Given the complexity of the values streaming from the sensor, an appropriate mapping technique was needed to extract useable control data. Rather than try to decipher this procedurally, a *black box* technique was chosen to deal with these complexities transparently. Echo State Networks (ESNs) proved valuable for solving this class of problem.

ESNs are a class of recurrent artificial neural network (RANN), which in turn are a special class of artificial neural network (ANN). ANNs have a unidirectional data flow, where a signal arrives at an input and propagates forward through one or more layers of neurons or nodes to reach the output. RANNs, however, have connections between nodes moving in any direction in the network, creating feedback loops in the flow of data. These feedback loops can create time delays, in effect giving a RANN short-term memory. RANNs are capable of modelling non-linear dynamical systems, making them ideal processors for temporal data streams, such as musical control data. They offer the possibility of modelling the complex web of convergent and divergent mappings observed in acoustic instruments. There are, however, practical problems with these networks; they exhibit complex behaviour relative to ANNs, so training them to perform a particular function can be difficult or unreliable. These issues were evident in the first type of RANN I trialled for mapping the sensor data, a continuous-time recurrent neural network (CTRNN) (Beer, 1995). This was trained using evolutionary techniques. From a user's point of view, the process of evolving these networks for mapping the foam sensor data could give inconsistent results, with each iteration potentially taking a long period of time. ESNs overcome these problems; training can last seconds instead of hours, and the process of training is simple in comparison.

ESNs are a fairly recent development in neural network research, belonging under the banner of *Reservoir Computing* techniques (Lukosevicius and Jaeger, 2009). At the heart of the reservoir computing concept is the realisation of computational power from non-linear dynamical systems. This is not necessarily limited to neural simulation, although this approach has been very successful (Verstraeten, 2009). Verstraeten lists other media including bacteria (Jones, Stekel, Rowe, and Fernando, 2007), a bucket of water used for speech recognition (Fernando and Sojakka, 2003), and potentially even the universe (Lloyd, 2002).

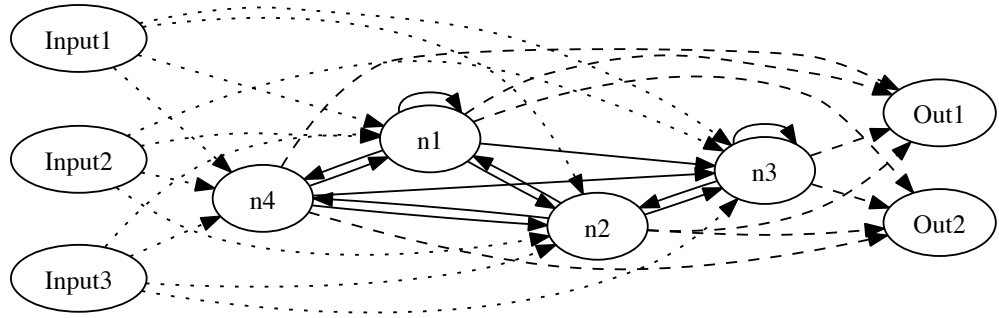


Figure 6.2: An Example ESN

Figure 6.2 shows a simplified example of ESN topology. The interconnected nodes n_i are the *reservoir*. A set of inputs is connected to the reservoir nodes, and these nodes are connected to one or more output nodes. The network is updated recursively as follows (Verstraeten, 2009, p. 32):

$$x[k+1] = f(\mathbf{W}_{res}\mathbf{x}[k] + \mathbf{W}_{in}\mathbf{u}[k]) \quad (6.1)$$

$x[k]$ is the network state at the current time step, $\mathbf{u}[k]$ is the current input matrix, \mathbf{W}_{in} is the matrix of input weights and \mathbf{W}_{res} denotes the reservoir weights. f is a smoothing function, commonly either a linear mapping, \tanh or the fermi function, $f(x) = 1/(1 + \exp(-x))$.

The key to the success of ESNs is the method of training; only the output weights are modified during this process. All other weights, in the input matrix and the reservoir matrix, are initialised with a random constant. The output layer is adjusted to exploit the dynamics of the reservoir and achieve the desired behaviour. Only one layer of weights is being trained, so training is a relatively straightforward linear problem that can be solved quickly using linear regression.

To function effectively, ESNs should possess the Echo State Property (ESP), meaning that the network has a slowly fading memory of its inputs. The presence of the ESP is dependent on the spectral radius of the network, a global scaling factor of the reservoir weights. This variable controls the richness of the dynamics and the non-linear modelling power of the network, at a trade off with its memory capacity (Butcher, Verstraeten, Schrauwen, Day, and Haycock, 2010).

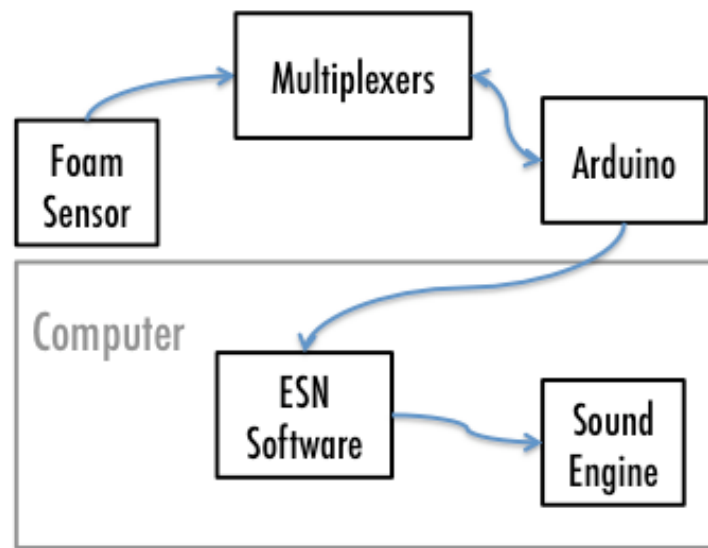


Figure 6.3: System Overview

ESNs have been used successfully in a variety of real-world applications, for example data-mining (Lin, Yang, and Song, 2008), industrial control (Qiao, 2009) and robotics (Ishu, van der Zant, Becanovic, and Ploger, 2004). They have also been applied to computer music problems. Holzmann (Holzmann, 2009) explored the use of ESNs for audio processing, using them for tube amplifier emulation and nonlinear audio prediction. In the context of processing sensor data, useful applications include sequence recognition and dimensionality reduction.

6.2.3 SOFTWARE IMPLEMENTATION

The ESN is implemented using Holzmann’s *Aureservoir* C++ library³. A program running in the OpenFrameworks⁴ environment controls the entire process, reading sensor data from the Arduino, mapping it through the ESN and sending the output streams to a sound engine running in SuperCollider (McCartney, 2002) via Open Sound Control messages. This program also provides data stream visualisations and enables control of the training process. Figure 6.3 illustrates the overall process.

6.2.4 ESN TRAINING

To train an ESN, a set of input data and corresponding output data needs to be created that defines how the trained network should behave. The control software facilitates the creation of training data, which is recorded in realtime; the foam is manipulated with one hand and computer number keys are held down with the other to set individual ESN outputs to zero or one. The complexity of training data required for good results is dependent on the behaviour desired of the system, and is best found through experimentation by choreographing varying sequences of inputs and corresponding outputs.

³<http://aureservoir.sourceforge.net/>

⁴<http://openframeworks.cc/>

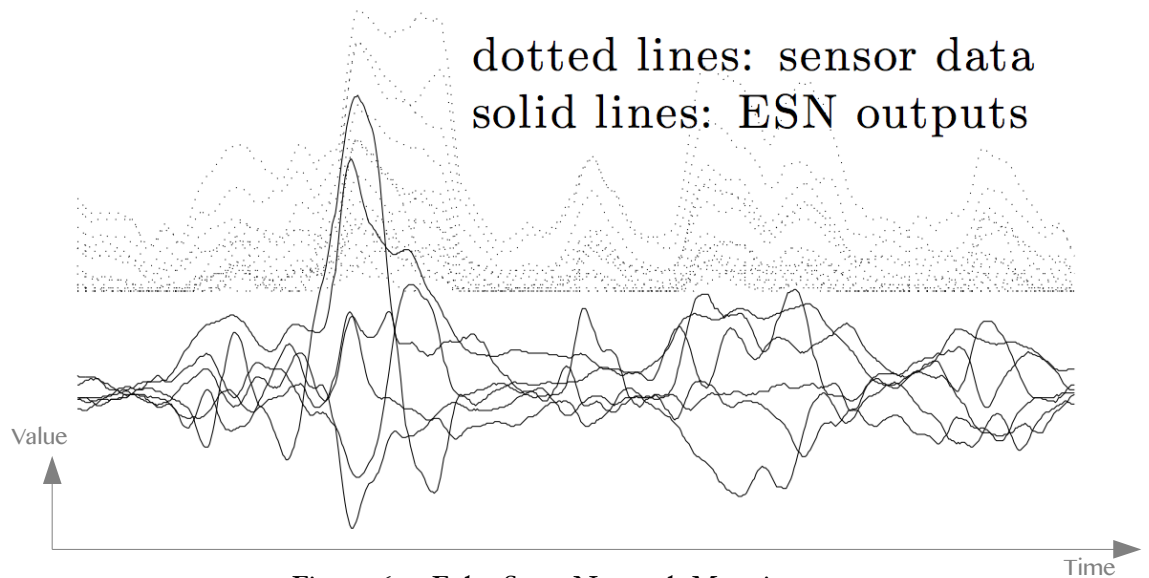


Figure 6.4: Echo State Network Mappings

6.3 Formative Evaluation

As with Phalanger, the system was evaluated in two stages, firstly with a formative, lab based study, and then subsequently with a field study. The formative study took place in early 2010.

Haptic exploration of sound synthesis parameter spaces was chosen as a context for the first evaluation of this system. After experimentation with ESN mappings, it was found that ESNs can be trained to output an arbitrary number of continuous data streams that change consistently with the position of the foam, essentially acting as a dimensionality reduction or expansion engine. A training set to achieve this behaviour is created by moving an area of the foam while holding a single output at a high value, and then repeating this for other areas and other outputs. Between moving these areas of the foam, the foam is left to settle while all outputs are kept at a low value. Figure 6.4 shows an example of an ESN trained in this manner; the 16 input streams are converted to six output streams.

The nature of the training process means that the output of the ESN after training is arbitrary, but nevertheless consistent and potentially musically interesting. These output streams are used as control data for sound synthesis patches; as the player manipulates the foam the sound changes in accordance with its position, allowing the player to explore the sound space with touch, gesture and physical manipulation.

The evaluation comprised a set of interview sessions, lasting approximately thirty minutes each, where the participants experimented with the controller in various scenarios designed to test the affordances of the controller. The participants were interviewed about their experience, the interviews being recorded as audio for later analysis. For all the scenarios, the foam controller was set up to be mapped to six continuous control streams through an ESN. The ESN was configured with 16 input nodes, 150 linear hidden nodes and 6 linear output nodes;



Figure 6.5: Kenton Control Freak MIDI Controller

the reservoir had a connectivity of 10% and a spectral radius of 0.8. It was trained with the pseudo-inverse algorithm, and the simulation ran with the *SimSquare* algorithm. Inputs to the ESN were mapped to between -10 and 10.

To provide a reference point, participants also tried controlling the same patches with six of the sliders on a Kenton Control Freak Studio Edition MIDI controller (figure 6.5); this represented a more conventional mode of control for sound synthesis.

6.3.1 SCENARIOS

The first scenario was the control of a phase modulation synthesis patch below (in SuperCollider code).

```
(
SynthDef(\pmsynth,
{
    |p1=100,p2=100,p3=1,p4=100,p5=100,p6=1|
    var w = PMOsc.ar(p1, p2, p3, PMOsc.ar(p4,p5,p6)).dup;
    Out.ar(0,w);
}).add
)
```

The *PMOsc* unit generator implements a phase modulation oscillator pair. The three arguments are for carrier frequency, modulation frequency and modulation index.

In Phase Modulation (PM) synthesis (Cavaliere, Evangelista, and Piccialli, 1988), tones are created by modulating the phase of a sine (*the carrier*) with another sine wave (*the modulator*). It was chosen as it is commonly regarded as being highly unintuitive to program, an interesting challenge for an interface that attempts to provide intuitive control. This patch provided a large, varied and non-linear timbre space for the participants to explore. The controller was mapped to work within the range of 20Hz to 20KHz for the carrier and modulator frequencies, and 0 to 100 for the modulation index.

The second scenario was a sound mixing task, where the six output streams controlled the volume of six looped variations of a vocal sample. These samples modulated together into a continuous soundscape that shifted subtly with the variation of amplitudes.

The final scenario was a variation on the first PM synthesis patch, where the range of control was constrained such that participants were working within a subset of the much larger parameter space. This gave finer control over the sound, providing a different experience from the first PM patch. Table 6.1 shows the ranges.

Table 6.1: Constrained Parameter Ranges for Scenario Three

Carrier Frequency 1	230.5Hz - 259.6Hz
Modulation Frequency 1	258.3Hz - 304.7Hz
Modulation Index 1	0.2 - 5.2
Carrier Frequency 2	172Hz - 4092.6Hz
Modulation Frequency 2	1374.3Hz - 1379Hz
Modulation Index 2	0.9 - 15

6.3.2 PARTICIPANTS

The participants were eight university students (two female, six male, aged between 18 and 40), all with experience in computer music. When asked to rate their skill in FM synthesis on a scale of one (none) to seven (expert), they responded with an average of 3.125, in a range from 1 to 6. They were also asked to list their total years of experience using separate computer music software packages, the sum of these years averaged 12 years, in a range from 5 to 19 years. Participants had an average of 5.86 years of musical training on consecutive instruments, ranging from 0 to 31 years. All participants were unpaid volunteers.

6.3.3 METHOD

In order to avoid influencing the participants' initial impressions of the controller, they were at first asked to begin exploring it without explanation of how it worked. After exploring scenarios one and two with both controllers, they were interviewed about their initial impressions, and the details about the workings of the system were then explained to them. They continued to try out the third scenario with both controllers and were interviewed again about their impressions of the system. Finally they completed a questionnaire on their preferences between the two devices in different contexts, and their responses to this were used as talking points for a final interview.

6.3.4 RESULTS

The eight interviews were transcribed and analysed using a grounded theory approach (see section 4.3.5.2). The participants' responses clustered around the following concepts:

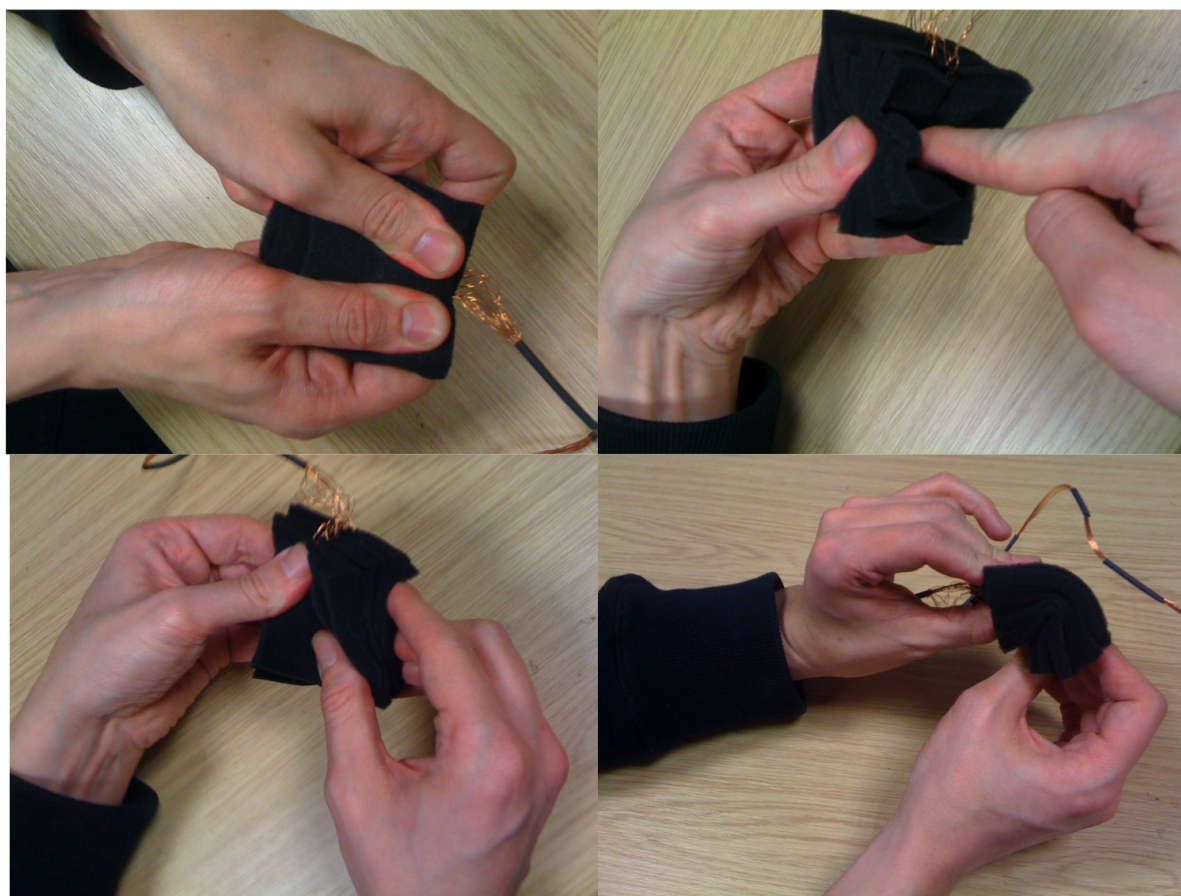


Figure 6.6: The Controller In Use

Intuitiveness Participants generally perceived the foam as a natural and organic method of control (*'I think [the foam] seems more organic, more natural, maybe intuitive, biological ... You weren't thinking about this region as much as this region, it was more about the tactile feel as opposed to looking'*). Some felt a more direct connection to the sound (*'It felt I was moving the music with my hands rather than moving controllers of it'*).

Learnability Participants perceived a range of learning curves, some finding the controller initially difficult to use (*'I couldn't get my head round when I was poking about with it, what that was doing'*), some finding it easier (*'you can just use it straight away, and practising with it gives a different sort of experience', 'manipulating sounds creatively, you can get some really interesting stuff out of that straight away, without prior planning'*). Most participants seemed to become more skilled with the foam over the course of the interview, although this may have been due to initial caution about breaking the sensor. Some thought it was instantly accessible (*'I think it would be really cool for people who don't really understand what's going on with MIDI to just play around with the music, and maybe kids'*) but would be difficult to master (*'I think you'd have to practice with it a long time to actually get the hang of it'*).

Physical Manipulation The freedom of movement with foam was something that some appreciated (*'The slider is two dimensional, foam is anywhere really, you can squeeze it from any angle, any pressure. This [slider] is just up and down', 'you were having a lot more control over sound because you could turn and twist'*). A wide range of motions were discussed, including squeezing, poking and twisting. Fingers, palms and hands were used to manipulate the foam, both single and dual handed. Some pressed the foam against the table, or squeezed one part of it while manipulating another.

High Dimensional Control It became clear in this sense that the foam was a markedly different approach to control from the sliders (*'I tried to establish the different parameters and really make them independent from each other, with foam it's very complex anyway because it's all entwined and there are no independent parameters, and all of a sudden it becomes a much more holistic abstract way of interacting', 'It's more related to human touch than it's related to the very limiting one dimensional moves that you can make with the MIDI controller.'*). This was an issue for some (*'It's a lot more difficult to find independent dimensions'*) but not for others (*'I felt like I had the whole music in my hands', 'With the MIDI controller I felt like I had control over one individual parameter at any time but it was quite difficult to control the whole shape of the music'*). The individual differences in preference to high dimensional control is a recurring theme, possibly relating to personality type or interaction style preferences. It is revisited in more detail in chapter 8.

Control Accuracy was an issue, in comparison to the sliders (*'You can get the levels of the sounds you like on the sliders quite easily, but it's a bit harder to get that exact sound that you're trying to achieve in foam', 'I find the foam at times a bit too coarse and a bit too crude'*).

Mapping Some participants felt the foam worked better with certain types of sounds (*'I can imagine using it one handed and it being a fun extra thing to use, for more effects based things', 'It actually felt like you could use it musically and have some control over it if you got used to it, especially on something where you were more controlling spectral stuff as opposed to pitch'*). Some enjoyed the correlation between motion and sound (*'[the foam] is interesting especially when you have these sounds where you squeeze something and you really feel that as you squeeze it the sound becomes tight and there is a correspondence between the actual physical activity of squeezing something and releasing it and also the sound became a bit tighter', 'It's very satisfying to express your relationship to what you hear ... if you hear something that's quite a hard sound you can also be hard about it in the way you touch the foam', 'It was kind of fun when you did properly manipulate the foam, you were crushing it and it did really go [makes crushing noise]'*).

Creativity One person felt that with the sliders, (*'you feel like you're operating a machine. I think [the foam] feels a bit more creative'*) while another felt that the precision of the sliders was more creative (*'You can control each little sound of the patch a bit more independently, I think it's a bit more creative in a sense'*). Some felt the foam was better for experimentation (*'It's about experimenting that you wouldn't make if you were looking for an exact sound. You're just stumbling across something, 'If it's more like doing weird stuff, then use the foam'*).

Visual References One participant requested to look at the data streams on screen as a reference, and commented on their experience (*'I think that what happens when you start to have a visual maybe that you become a bit more goal oriented, you are trying to move the sliders or move some sliders versus other sliders, which I didn't do before, it was purely just an internal experience', 'once you have a visual it becomes something else, because in a way your attention is divided'*). Another participant felt a visual reference would help with accuracy.

Repeatability There were mixed reactions about repeatability with the foam (*'[would you find it easy to go back to the same manipulation?] Yes because there's a body memory where your fingers kind of know where to go', 'It's easier to remember the way in which you twisted a piece of foam than it is to remember the positions of six different sliders, so it meant that I could, with a bit of practice, go back and forth between different settings', 'Sometimes I was trying to make the same thing happen twice but I couldn't always make that happen', 'There was one point where I had absolutely no idea how I'd made a sound and I didn't know how to get it back', 'You're not going to be able to necessarily remember that particular shape that you twisted it into to be able to use it again'*).

Exploration Participants were asked which interface they preferred for exploration; some preferred the MIDI sliders (*'I had more control over the different parameters so I could keep the rest constant and vary the others, whereas I didn't know how to control the parameters, I couldn't isolate one with the foam', 'You can turn everything one by one and learn what the components are and then you build up the sound enough'*), while others chose the foam (*'It was more fun just to explore around in the sound space and it meant that I could keep playing with it and find places where it sounded interesting', 'I guess as soon as you see six faders ... I was kind of methodically going through them ... that's really working out exactly what's going on, but that isn't necessarily exploring the soundspace in an interesting way, it's a very sort of stepped, obvious way, it's a lot more interesting to be doing it with the foam. Because you get results that maybe you wouldn't have done with just moving faders up', 'It's a much more intuitive approach because even though you get an idea where things are, there's so much mapping going on that you don't have a clear image in your head of where they lie so you explore it in an intuitive way and combine them in an intuitive way .. I can play with this a lot longer because it seems like there are more combinations I can make of bringing the sounds together than what I can do with the more structural MIDI controller'*).

Applications Several participants commented on the foam's potential as a collaborative tool. There were also comments about its potential as a performance tool, some negative (*'I'd like to have more control in performance'*), some positive (*'it fits more into the live set of music creation rather than sitting in a studio'*), and some from an audience perspective (*'It would visually be interesting', 'I think for performance, watching someone making motions that are in line with the human body in a sense that they are fluid is more aesthetically pleasing than seeing somebody moving a slider'*).

Fun This was a prominent theme in the responses about the foam (*'I had much less control ... over what was happening but that kind of made the foam thing more fun', 'It's just always a lot more fun when you're just using your hands in a natural way', 'This one's more fun to play with, more engaging without a doubt'*).

6.3.5 DISCUSSION

Having presented a summary of the results, before discussing them further it's also necessary to consider their validity in the context of the study that was performed. One shortcoming of this short interview form of evaluation is that there's a novelty factor that may influence the responses. Two participants also commented that they disliked the patches they were playing (*I found it harder to understand how I was controlling it, maybe I just didn't like the sounds*), which may also have affected their opinion of the controller. These problems would be ironed out in a field study, which would give participants time and space to practice with and find the creative limitations of the controller on their own terms. What these results do provide is a useful set of pointers from which a picture of the controller can be built; its strengths, weaknesses, and outlines of the themes and issues that concern and influence its design.

A notable strength of the system was its intuitive feel, participants describing it as feeling natural and organic, and giving them a direct tactile connection with the sound. The controller affords a freedom of motion in interaction, and a wide ranging vocabulary of manipulations were tried during the study. An interesting aspect of this is sound-motion correspondence, where the sound being generated is perceived as correlating with the physical manipulation of the controller, for example squeezing the foam causes the sound to become *compressed*. While a slider that can be moved up and down may only correlate with rising or falling elements of a sound, a malleable controller with many more degrees of freedom has potential for a wider range of correlations; mappings could be designed to deliberately exploit this feature.

An obvious weakness of the controller was accuracy. There are two factors in the design of the system contributing to this; firstly there is a small instability in the output of the ESN such that the output streams oscillate slightly. This is more noticeable in some patches than others. The nature of the foam itself is the other factor; when compressed and released there's a period of expansion where it returns to its original form, so there is inherent motion in the output. This issue of accuracy is closely tied in with the issue of repeatability, a topic on which participants gave a mixture of responses. The controller relies on a mixture of visual, tactile and proprioceptive senses for precise control, so it could be regarded as quite difficult to use accurately. It also relies on a vocabulary of gestures that are less commonly used for musical control; it's unclear from this study how practice might improve these issues, a subsequent field study shed some light on this. The style of mappings that some participants stated a preference for follows on from these issues. They perceived the controller as being more useful for settings which required a lower degree of precision.

The high-level manner of control was a prominent issue. With the interdependent nature of the parameters, and the nature of malleable control where whole motions correspond to changes in sound, the underlying synthesis mechanism is obscured and the foam becomes an abstraction of this mechanism, the sound becoming embodied in the controller. Some participants saw this as a strength of the controller, reducing the cognitive load of engaging with the underlying mechanism and promoting a fluid style of interaction. Others found this awkward; they naturally approached the foam as they would the mixer, attempting to separate dimensions, though in an interface with a much larger freedom of control these dimensions are more numerous and less separable.

In terms of the creative potential of the controller, some felt that this was increased by its more imprecise nature, while others felt this detracted from it. (*I personally feel that if I'm going to be creative I like to be a bit precise*). This issue is closely tied in with the theme of the controller's intended use in the evaluation scenarios, as a mechanism for the exploration of sound spaces. Again the controller elicited a range of responses on this topic, participants who took a more intuitive, unpredictable approach to exploration tended towards expressing a more positive reaction to the controller than those who worked in a more methodical manner. As Gelineck and Serafin (2009b) observe, musicians seem to like a tool that has 'a life of its own', so unpredictability and imprecision can have a useful place in the composition process.

6.4 Design Improvements

Following on from feedback in the formative evaluation, several improvements were made to the design of the controller. Feedback from a demonstration at the New Interfaces for Musical Expression conference in 2010 also contributed to these changes. These improvements solved some functional problems and also aimed to improve the experience of using the controller. The primary aim was to make the EchoFoam controller ready for evaluation 'in the wild'.

6.4.1 PACKAGING

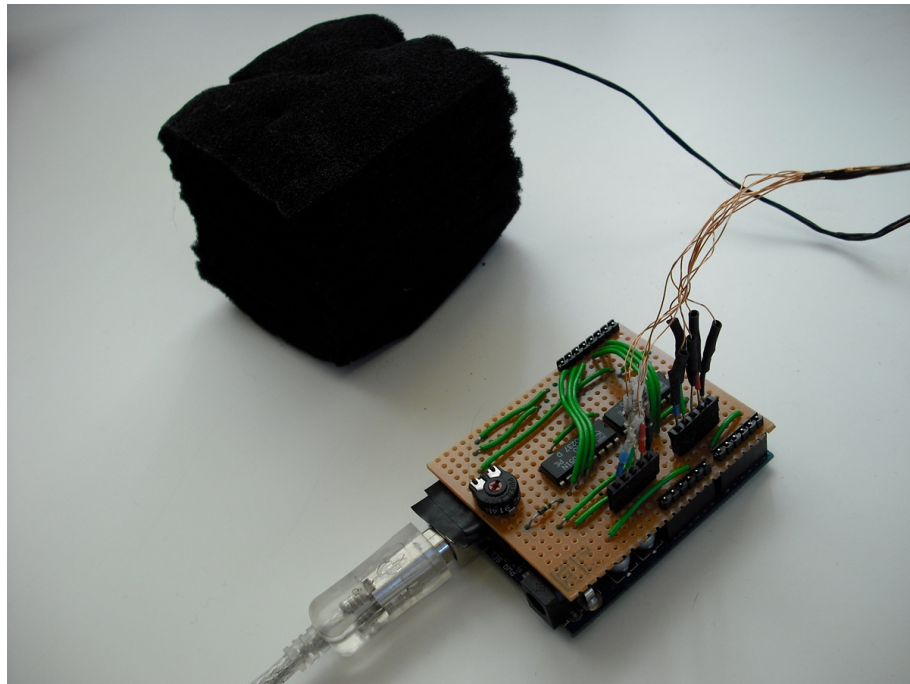


Figure 6.7: An EchoFoam Controller, Version 2

For the formative evaluation, the circuitry for the controller was connected together on a breadboard. This delicate setup needed to be changed into something robust that could survive real-world use. To help achieve this, the circuit was made into an Arduino shield (pictured in figure 6.8). This is in essence a circuit board that can plug directly in to an Arduino.

Making a PCB was outside the scope and budget of the project so the shields were constructed using stripboard with pin headers soldered in to make the connection with the Arduino. Improvements were also made to the set of wires between the foam and the circuit; instead of being taped together, they were merged into a single cable using ‘liquid tape’, avoiding loose wires. The new version of the controller is shown in figure 6.7.

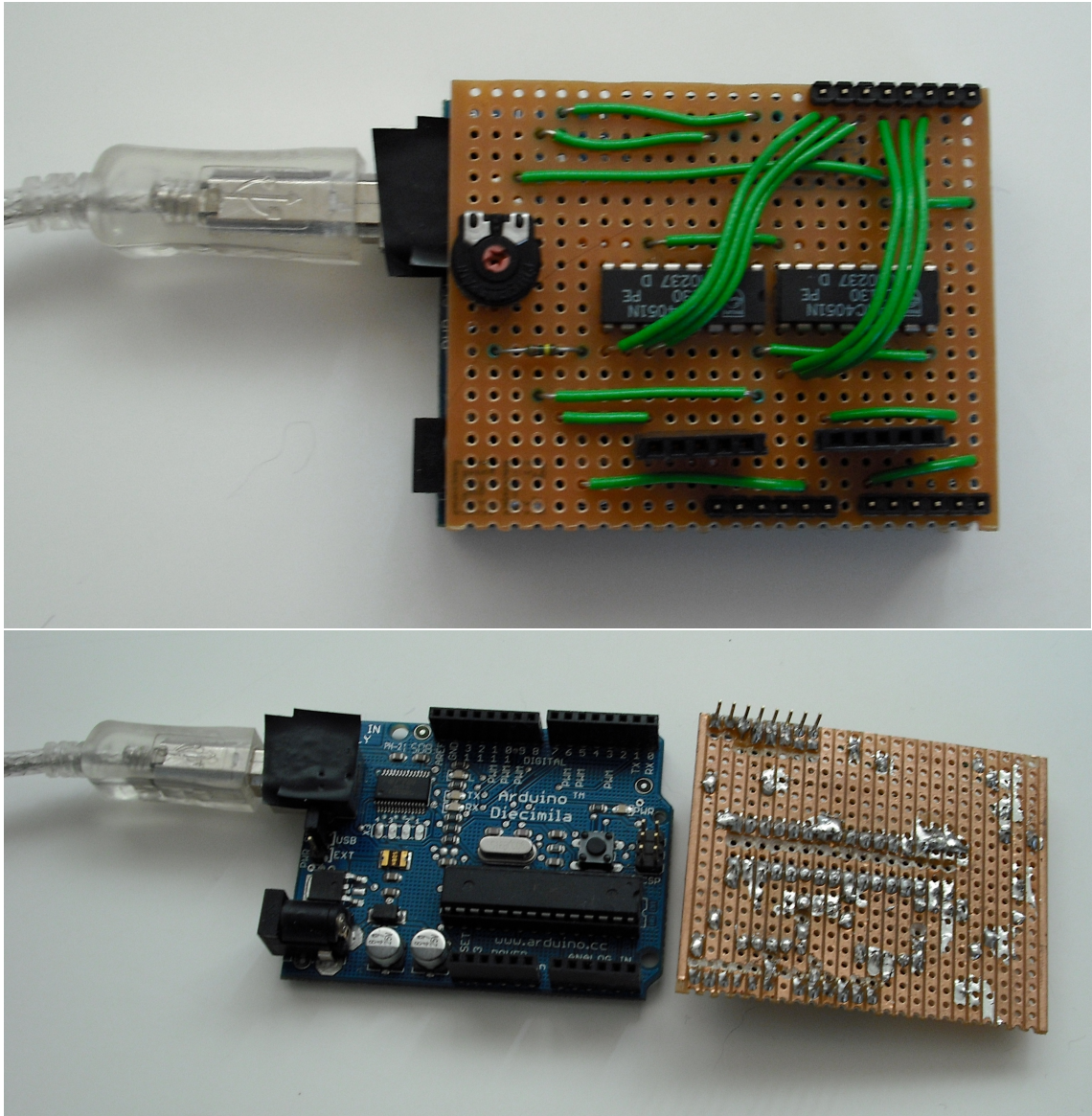


Figure 6.8: EchoFoam Arduino Shield

6.4.2 INCREASING CONTROLLER RESPONSIVENESS

In the prototype, the points of contact for the wires within the foam were placed arbitrarily over the area of the sensor. While this approach was successful, it could also be improved on. It was desirable for distribution of sensitivity within the foam to be more even, and also for the sensitivity to be increased globally. To achieve this, a new approach to measuring the resistance in the foam was devised, using two-way multiplexing. This approach achieves

increased sensitivity while simplifying the circuitry.

Resistance in the prototype controller was measured by sending a current through a single live wire and then measuring the resistance at 16 different contact points. The new design uses two sets of wires, one set of output wires and one set of input wires. Each set is attached to a different multiplexer, and code on the Arduino sequences measurements as follows:

```
foreach output wire outWire do
    pass a current through outWire;
    foreach input wire inWire do
        read the voltage on inWire;
    end
end
```

Figure 6.9 illustrates this process. Using multiplexing in this way, the numbers of resistance measurements taken from the foam can be increased with fewer contact points. In the prototype, $c = m + 1$, where m is the number of measurements and c is the number of contacts. With the new scheme, $c = 2\sqrt{m}$, assuming equal numbers of inputs and outputs.⁵

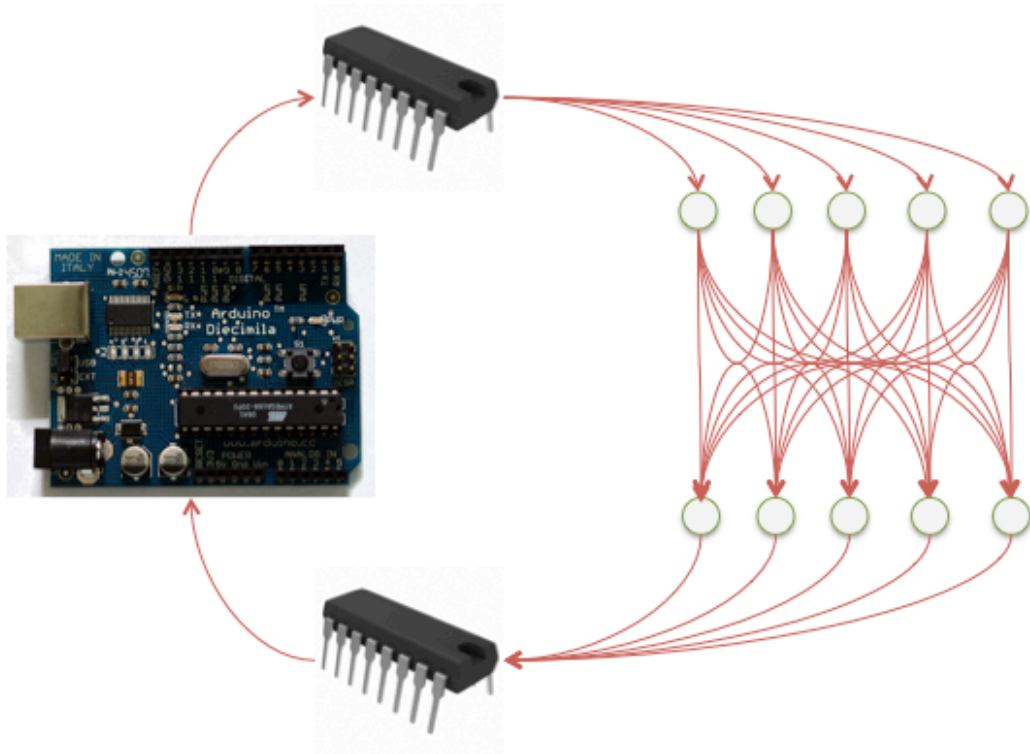


Figure 6.9: Multiplexing in the EchoFoam Controller

For the new version, a configuration of ten contacts provided twenty-five resistance measurements was chosen. This yielded 56% more measurements than the prototype. While more contacts could have been used, the number of readings needing to be processed would have slowed the whole system down too much, both at the level of reading measurements with

⁵The number of measurements is equal to the number of output contacts multiplied by the number of input contacts. For example, to create 16 measurements, four input contacts and four output contacts are needed.

the Arduino, and also at the stage of processing through the ESN. Using five contacts in each set also allowed even placement of the wires; the sets of contacts were tied into separate foam squares in the four corners and the centre (see figure 6.10).

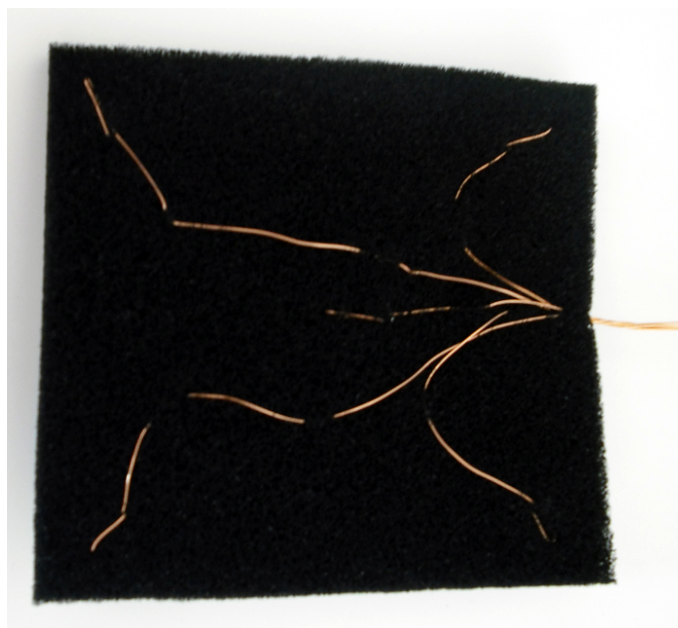


Figure 6.10: Enamelled Copper Wire Tied into a Section of the Controller

These pieces were then used as the top and bottom sections of the cube, and other pieces were glued together in between them to create the whole sensor (figure 6.11). In this version, a total of 12 pieces were used, giving a 7.5 centimetre cube.

Figure 6.12 shows the lines of measurement between contacts, with the foam in an undeformed state.

6.4.3 FOAM SENSITIVITY DEGRADATION

It was found that the foam loses sensitivity over time, as it is manipulated more and more by the player. This results in the voltage readings decreasing to a low level, so that the foam needs to be compressed with increasing force to achieve the same effect. To compensate for this issue, a preset potentiometer was added into the circuit (pictured next to the USB socket in figure 6.8); this could be used to make a global adjustment to the sensitivity of the controller. This was done at the level of analogue electronics rather than digitally in order to preserve a high resolution.

6.5 Field Evaluation

The next step was to undertake a field study, to further evaluate the controller. This study was carried out over the summer of 2010, when EchoFoam controllers were given out to computer musicians who used it in their own environment and reported back on their experiences. An application, entitled *EchoFoam Live Link*, was created as a tool to use the controller with Able-



Figure 6.11: The Foam Sections of the Controller, Before Being Glued Together

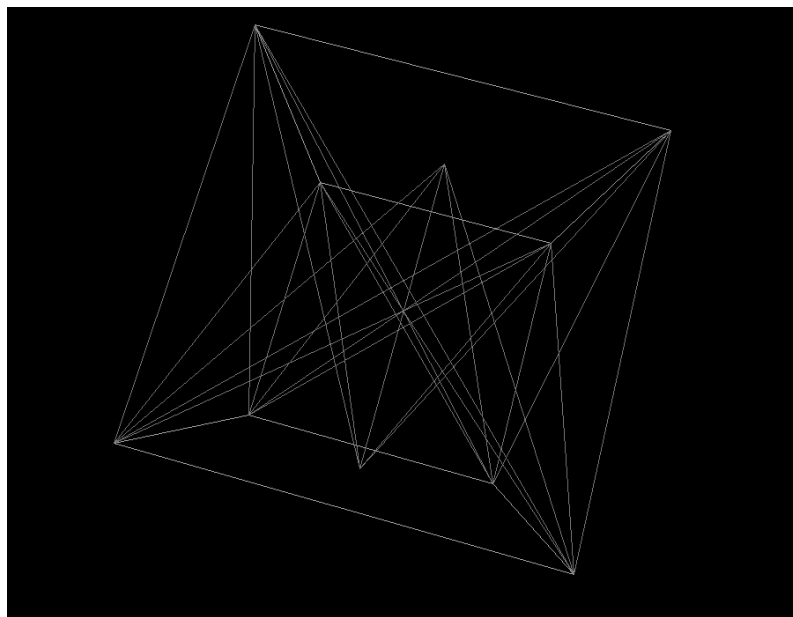


Figure 6.12: Lines of Measurement in an EchoFoam Cube

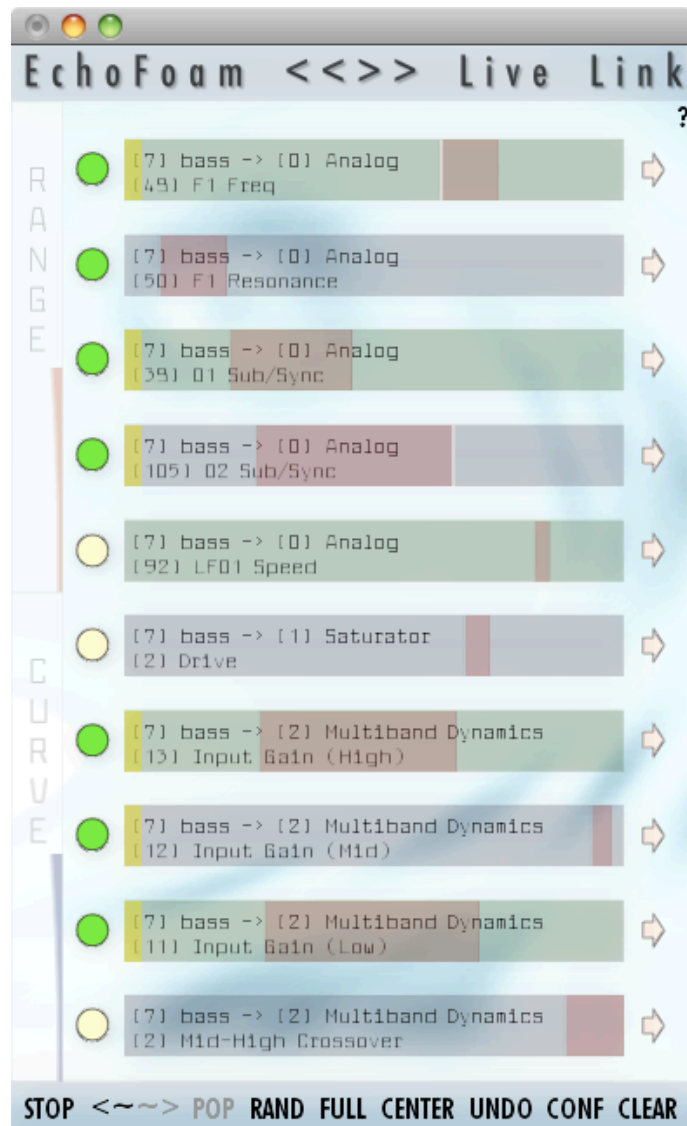


Figure 6.13: EchoFoam Live Link Screenshot 1

ton Live. I will start by describing this new software, before outlining the study and presenting the results.

6.5.1 SOFTWARE

With the focus of this thesis on interaction with digital music tools, I wanted to see how the EchoFoam controller would fit into a typical electronic musician's studio. It was decided to create software to accompany the controller that would integrate it into commonly used DAW software, making an entry point for the controller into musicians' everyday work processes. There were a number of options for this, the most popular applications including Steinberg Cubase, Apple Logic and Ableton Live. Ableton Live was chosen, principally because of the integration possibilities offered by LiveAPI, an unofficial Python scripting interface that sits within Live. The compositional approach offered by the software was another factor, the software being specifically aimed towards creation of sound and music as well as production

work. This provided a wide range of potential usage scenarios for the controller.

It was decided to include features from the timbre space search system, *RECZ*, described in chapter 7; the evaluation study showed that this enhanced the functionality of the controller, and this study provides an ideal opportunity to further evaluate these methods. A subset of features of *RECZ* were chosen that were suited to use with new software.

With *RECZ* included, the software was designed as a tool what would connect the EchoFoam controller to continuous parameters within Ableton Live, with *RECZ* used as a configurable and interactive mapping engine. This would create a fairly open space of possibilities for using the controller with Live.

6.5.1.1 IMPLEMENTATION

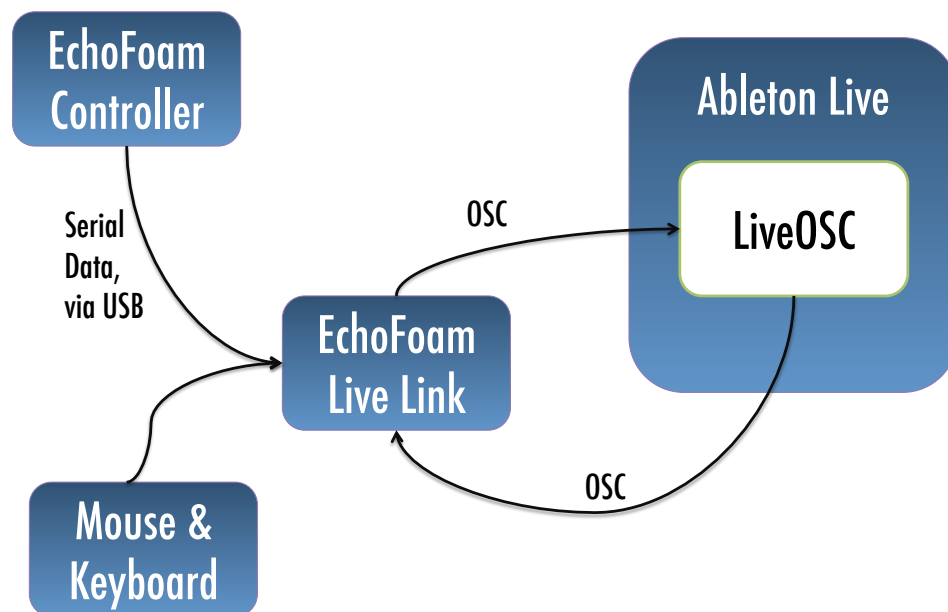


Figure 6.14: EchoFoam Live Link Data Flow

6.5.1.1.1 LiveOSC Ableton Live has an internal Python⁶ scripting engine to allow third party manufacturers to tightly integrate their controllers with it. It exposes an API that allows the querying of the state of most internal elements in Live, and allows control of parameters and settings. This scripting API is not officially supported by Ableton for use by the public, however it has been opened up to developers with the *LiveOSC* project.⁷ LiveOSC is an OSC server that sits within Live, allowing third party applications to query and control the Python API features with OSC commands.

⁶ <http://www.python.org/>

⁷ <http://livecontrol.q3f.org/ableton-liveapi/liveosc/>

To link EchoFoam Live Link (or EFLL) to Live, a modified version of LiveOSC was created, called *LiveOSCEchoFoam*. This added in some extra low-level functionality that was missing from LiveOSC. Figure 6.14 demonstrates how these elements fit together; the controller sends data to EFLL, which then communicates with LiveOSCEchoFoam to query the available parameters and set new values for them.

6.5.1.1.2 Tools The software was based on the software already put together for managing the EchoFoam controller. This was built using OpenFrameworks C++ library, and used the Aureservoir ESN library. OpenFrameworks provided an OSC interface for communication with LiveOSC. The software ran on the Apple Mac OS X platform.

Table 6.2: Echo State Network Configuration

Inputs	25
Outputs	1 .. 10
Reservoir Nodes	100
Connectivity	10%
Alpha	0.1
Reservoir Activation Function	Linear
Output Activation Function	Linear
Training Algorithm	Least square
Simulation Algorithm	Squared state updates

6.5.1.1.3 Echo State Network Mapping At the core of the mapping engine was an ESN used for dimensionality reduction, as described in section 6.3. EFLL allowed the use of between one and ten output streams, so the program had ten pre-trained ESNs ready to map the data, one for each number of output channels. These ESNs have 25 inputs, and outputs corresponding to the number of channels the ESN was created for use with. Table 6.2 shows the full configuration details.

To train these ESNs, a system was created for the systematic training of all ten networks. In a configuration mode in EFLL, configuration data was recorded of the foam being manipulated in a particular style of motion for each potential output. Data was also recorded for the controller at rest, and expanding to the resting point. This configuration data was recorded in separate pieces. When the pieces of training data had been recorded, EFLL created custom training sets for training the ten individual ESNs and trained them in a batch process. These trained networks were then stored. When the program was running, they were all held in memory so the program could switch between them as required with no break in output.

For the study, the ESNs were pre-trained for each participant's own controller.⁸ These networks were stored in a configuration file, installed along with their software. This freed them from the process of training, which takes some experience in itself to obtain good results.

⁸This separate training was necessary due to small differences between controllers. They were hand made, so there were subtle variations in contact placement, glueing and foam size.

6.5.1.2 FEATURES

Figures 6.13 and 6.15 show screenshots of the software. As part of the study, a website was created with installation instructions and a reference manual; the content from this is reproduced in appendix B. The salient features of the software will be outlined here, for more detail please refer to the manual.

The user could control continuous parameters in Ableton Live with up to ten channels in EFLL. These parameters could be selected in varying ways, and each channel had a number of mapping options. There were also features taken from RECZ which affected channels globally.

6.5.1.2.1 Parameter Selection To select a parameter, the user armed the channel in EFLL using the GUI, and then moved the parameter in Live. This parameter would be picked up by EFLL when it moved, and would now be under the control of the EchoFoam controller. The parameters could also be selected randomly as part of an explorative process; hitting the *Populate* button filled the selected channels in EFLL with randomly selected parameters from the current device in Live.

6.5.1.2.2 Search Features Each channel had a modifiable working range. The user could either change this manually with the GUI, or clicking the *Random* button applied random ranges to the selected channels. These ranges could be scaled using the range slider, and the ranges could also be centred on the current values. This allowed exploration of the parameter space similar to the previous EchoFoam study. A *Curve* slider changed the mapping for random range sizes between linear and sigmoid shaped, allowing the user to change the likelihood of exploring narrower or fuller spaces.

6.5.1.2.3 User Interface Features Each channel could be enabled or disabled, allowing the user to determine whether it sent control data. This feature was implemented so that the user could freeze parameters they no longer wished to control, with as part of the search process. Each channel could be *selected* or not, meaning that the user could control which subset of channels could be randomised or randomly populated. This feature set allowed global control of the search process together with fine tuning should it have been required.

Most features were controllable from the keyboard as well as the mouse. A debug mode (see figure 6.15) showed graphs of the live data streaming from the controller, and the mapped data streams from the ESN that were sent out to Live.

6.5.2 STUDY DESIGN

6.5.2.1 PARTICIPANTS

After obtaining ethics clearance, participants were advertised for on internal University of Sussex email lists, and this email was passed on by individuals to other interested people outside of the university. The email asked for Ableton Live 8 users with Apple computers who were interested in trying out the controller. The participants would get to keep the software and controller as payment for taking part in the study.

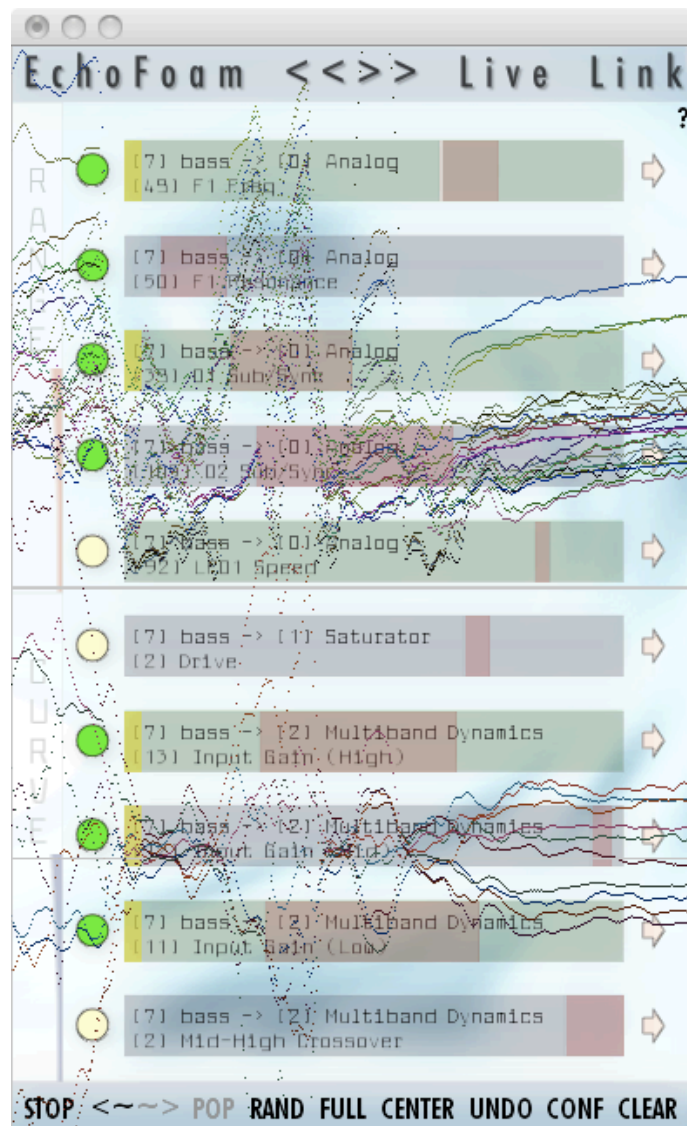


Figure 6.15: EchoFoam Live Link Screenshot 2: Showing Data Streams

The call attracted thirteen participants, from which the first ten were invited to take part as there was only time and resources to make ten new controllers. Half of the participants were located locally and the other half were from other places in the UK. For a variety of reasons, four of the participants could not complete the study. Of the six who completed, one was female and five were male, and their ages were between 18 and 40.

After the final interview, participants were asked to complete a background questionnaire. Five of the six participants returned responses. To give a picture of those who took part, here are the questions from the questionnaire and a summary of their responses. For questions 1 and 2, the number of years of experience for each person was calculated as the sum of the years of experience with each instrument or tool.

1. *Please list any acoustic instruments that you play, along with the number of years of tuition you have had with each instrument, and the number of years that you have actively played it for.*

One participant had no experience of acoustic instruments; the others all played two or three instruments including drums, piano, guitar and voice. The average number of years of active play was 21.2 (min: 0, max: 46). The average number of years of tuition was 6.5 (min: 0, max: 24).

2. *Please list the main computer music software applications you have used, and the number of years of experience you have of actively using each package.*

The following software had been used by the participants: MAX/MSP, Ableton Live, Traktor Scratch, Pro Tools, Logic, Cubase and Common Lisp Music.

The average number of years of experience was 11.3 (min: 10, max: 13).

3. *Please list the musical controllers you commonly use (MIDI or other).*

The following controllers were listed: Kenton Control Freak Live, Monome 40h, Non-MIDI synths and effects units, Novation Remote 25, Novation Launchpad, Korg Nano Control, Behringer DDM 4000, Korg Microkorg, M-Audio Oxygen8, Generic joypad, 88 key weighted piano, Livid Ohm 64, Behringer BCR 2000 and iPod Touch.

4. *Please rate your expertise in Ableton Live on a scale from 1 to 7, where 1 means novice and 7 means expert.*

The average rating was 3.9 (min: 1, max: 5.5).

5. *Please estimate the number of hours you spent using the controller (just give your best guess).*

The average was 14.3 hours (min: 6.5, max: 25).

6. *Please estimate the frequency of use (e.g. once per week).*

Participants responded as follows:

- Two or three sessions over a couple of weeks.
- It was more like a long-ish session sporadically.
- Two per week.
- Once per fortnight.
- Usually around 3 times a week.

7. *Do you think you will continue to use the controller in future projects?*

All participants said that they would use the controller in the future.

6.5.2.2 METHOD

The study began with an interview and tutorial for each participant. The local participants were met with in person, while remote participants were sent the controller and contacted over Skype. During the interview, the software was set up on each participant's computer and they were given a tutorial on how to use the system. Further information was always available for reference using the website created for the study. Participants were told that they could get in contact with queries at any time during the study. They were also asked to record their experience in a way which suited them; for example with a diary, screen shots, audio examples or videos. These items would be brought to the final interview where they would be discussed. At the end of the period of the study, each participant was interviewed either in person or remotely using Skype. These interviews were semi-structured; there was a list of topics which the participants were asked to discuss or comment upon, but otherwise the conversation was allowed to take any direction. During in-person interviews, a controller was available to aid demonstration of motions and to aid participant recall.

6.5.3 DATA

The vast majority of the data collected in this study came from the interviews which took place at the end of the study. The interviews lasted on average for 44 minutes. Three hours and 41 minutes of interviews were recorded and transcribed, resulting in approximately 18,000 words of text. This was coded and analysed using a grounded theory approach (as described in section 4.3.5.2).

Only two participants submitted any media supplemental to the end of study interview. One person supplied some screen shots taken when they were using the controller and these were discussed in the interview. Another submitted screenshots, a brief audio file and a document they had used to take notes about their experience during the study. This document was coded and included in the grounded theory analysis.

6.5.4 RESULTS

Analysis of the results led to the discovery of five main categories: the controller, aspects of control, software use, mapping and the system in use. Supporting quotes are given where they supply further illustration.

6.5.4.1 THE CONTROLLER

Controller Design. Participants were asked about their feelings on the design of the controller. Most comments concerned the form of the foam cube. Three participants commented that the cube format was considered to work well.⁹ The controller could have benefited from being larger, particularly with performance in mind.¹⁰ Alternative shapes

⁹Participant 4: 'I think it's a brilliant shape because it's got all these different approaches', participant 3: 'Generally I think the square works quite well. There's points on it for reference, that's why you can remember where things were which is handy', participant 5: 'And I like the idea of the cube, it's kind of rubiks, you're kind of used to that sort of thing'

¹⁰Participant 1: 'I don't know, maybe just because it felt a bit... like if you were standing holding the cube you feel a bit naked or something, there's not much there.'

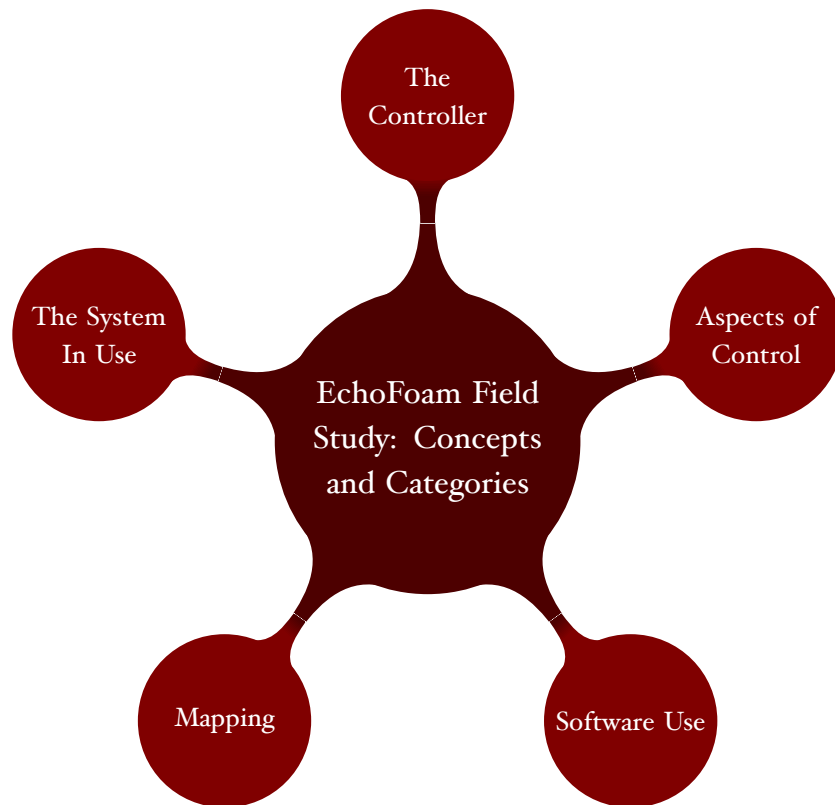


Figure 6.16: EchoFoam Field Evaluation Interview Analysis: Main Categories

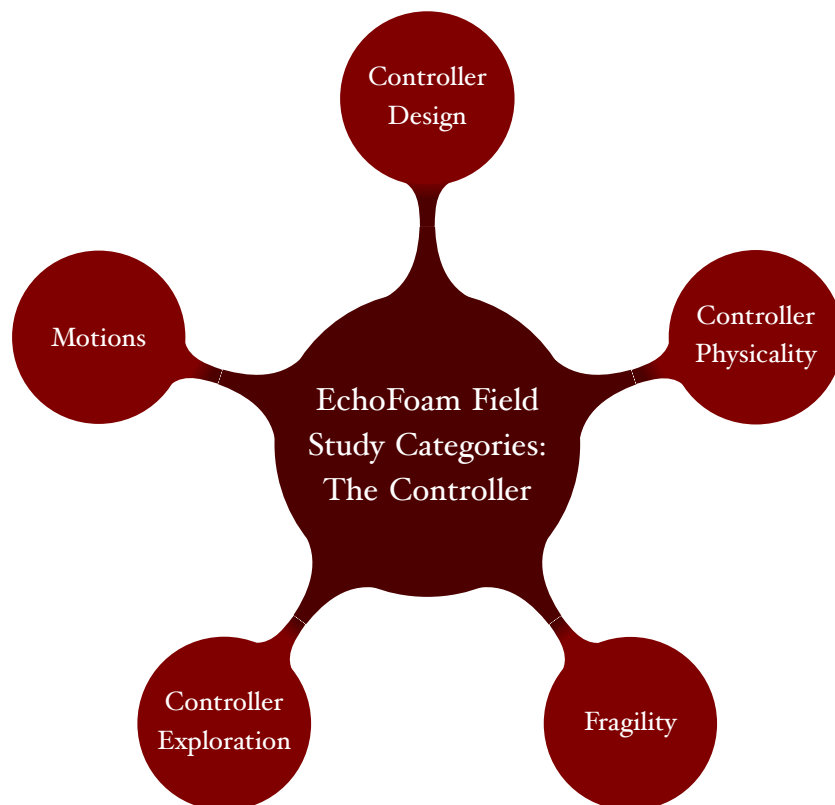


Figure 6.17: EchoFoam Field Evaluation Interview Analysis: The Controller

were suggested such as a pyramid, a ball and a flat panel. One suggestion was to make completely customised shapes, for example creating a controller based on an artist's logo. Another suggestion was to make a set of varied shaped controllers which could be used in different scenarios as required.¹¹ Some participants considered the possibilities of different foam densities, one feeling that denser, slower foam could be better.¹² Participant five felt constrained by the wires, proposing that a wireless controller would be an improvement.

Controller Physicality. Two participants commented that they liked the tactile feeling of the controller¹³, although participant three felt it was not satisfying enough to use¹⁴ because of the density and texture.

Fragility. Three participants perceived the controller as fragile when they first used it, making them approach the controller at first with caution for fear of breaking it.¹⁵ This caution generally evaporated after they had explored the controller, and all the controllers survived the study without any breakages. Other participants did not feel this initial caution and were happy to explore the device without hindrance from the beginning.¹⁶

Controller Exploration. There were varying approaches to exploring the ways in which the controller could manipulate sound. A common theme was to try and separate the cube into different areas, and see what each one would do with the current mapping. The participants divided the cube geometrically¹⁷ into corners, sides and areas and manipulated these. After this initial approach, participant three talked of how they started to explore the cube with whole motions, trying different motions such as squeezing, twisting, flattening.¹⁸

Motions. Participants were asked about the types of motions they used to manipulate the foam controller. Two participants talked of how they took a new approach with every

¹¹Participant 3: 'I guess there's no reason why you couldn't have a few of them linked into the same thing, perhaps different shapes and sizes, a set that you could unplug, and re-plug each one in depending on what you want.'

¹²Participant 2: 'it springs back to the original position quite quickly and I was wondering if the different more denser phone, what it would be like? Like ear plug foam. The reason I was thinking this was because I was thinking of it in a performance aspect, using it with the other devices maybe. Even though it's quite easy to use, it's kind of like squeeze it and then I'd be doing something else and then it finishes what it's doing quite quickly.'

¹³Participant 5: 'That you can really squidge it about, that it's a nice tactile kind of the thing', participant 1: 'It's just quite nice to have something soft ... the different feel of it is certainly a good thing I think'

¹⁴Participant 3: 'The only way you could improve it for me would be to make the whole physical squeezing thing feel more satisfying, perhaps different density of textures and stuff'

¹⁵Participant 1: 'It felt... seeing the wires for me made me feel a bit worried but it never would break but just sort of... in that connection way it felt like I had to be delicate with how I was picking it up, but when you were squeezing it was absolutely fine', participant 3: 'I think that the fact that I did get braver with it says quite a lot, because that was something that I did learn through the time using it. I learnt that it wouldn't come to pieces if I really was rough with it.', participant 6: 'When I first started using it I thought it was more delicate than it was so I was a bit more gentle on it, but as time went on I started to get a bit more rough with it, and then that seemed to make some better sounds.'

¹⁶Participant 2: '*Did you perceive it as being fragile?* Not really, I see it as I could destroy it if I wanted to but I've given it quite a go, I've really squished it to pieces and it sticks together pretty well. I think because you can visually see the layers, you know it's been put together, then you know they could be taken apart I guess. It still holding together pretty well but I can lift it up a bit at the corners and stuff. But I've put it through its paces', participant 1: 'it was very durable ... you could tell it wasn't going to break'

¹⁷Participant 4: 'I never really saw it as just one controller, that's the thing. I just saw it as the right corner and the left corner, as separate functions', participant 5: 'I went through the whole thing of just trying to do all the corners and ... I was trying to break it down once I had set up a load of stuff'

¹⁸Participant 3: 'I think at the start it was more about squeezing the corners, squeezing it from the top of the shape and pushing the top and the bottom of it together, whereas afterwards I started to actually just squeeze the whole block, morph it into different positions. In the end I wouldn't be keeping it as a square shape at all. The sounds seemed... it was more extreme when you treated it like that.'

new session and sound, much depending on what the cube was mapped to in Live and on how the mappings were configured.¹⁹ Three participants favoured particular motions, mainly squeezing, but also twisting and crushing the controller.²⁰ Two people discussed how they used very fine-grained motions to control sound, talking of sensitive *hotspots* in the foam where detailed manipulations could be made for a particular patch.²¹ Participant 4 discussed how one of their collaborators used the controller rhythmically, punching and squeezing different points repeatedly.²²

6.5.4.2 ASPECTS OF CONTROL

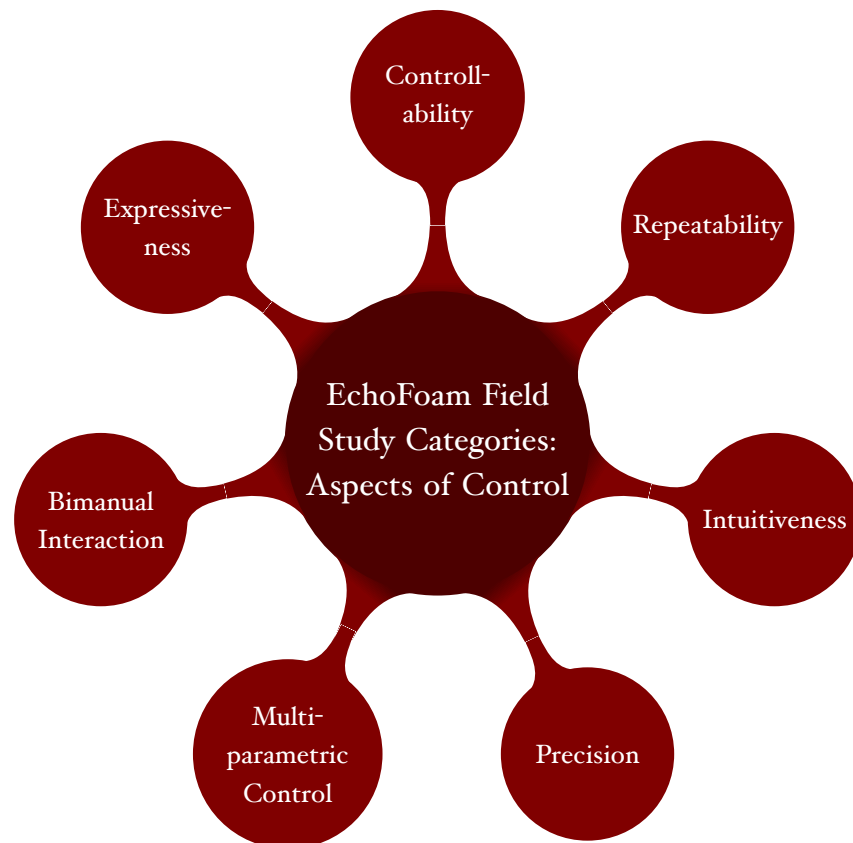


Figure 6.18: EchoFoam Field Evaluation Interview Analysis: Aspects of Control

Controllability. Discussions centred around the lack of control participants felt they had with the controller. Two participants felt that it was just too unpredictable, while others found a mixture of unpredictable and predictable elements, some learning to adjust to

¹⁹Participant 2: 'In a particular session I would get to certain motions for that because that particular channel was mapped to a certain control in the device that was doing the most satisfying sounds at that time. That would change every session, I would put different things in different channels'

²⁰Participant 2: 'Squeezing more than anything else, quite often I would have it on the desk and just squash it down and like rotating more pressure and then maybe using a finger to scrunch it,'

²¹Participant 4: 'I was approaching it by just squeezing and bending slowly then trying to find the little hot spots, the tickle points', participant 1: 'and then say the sound I wanted was like there, but I'd have to have the thing squeezed, and then my slight hand movements are going to open that up a bit and change the sound'

²²participant 4: 'he was rhythmically pushing it but kind of remembering roughly how he was pushing it and developing little patterns rather than just kind of doing slower squeezes and bends which is what I'd been doing up to then. He started rhythmically punching it and squeezing it one at certain spots again and again. That was actually a really good approach I would say.'

or embrace the less controllable elements of the system.²³ Of the users that felt it was too uncontrollable, they perceived the controller as working randomly²⁴ and found this frustrating.²⁵ Lack of control was considered by two participants to be both a strength and a weakness of the system.²⁶ The degree of controllability was heavily linked not just to the physical controller but to the system as a whole; the mappings to Live and the synthesis modules being mapped to. As the number of channels mapped to Live increased, so did the instability of the system. This instability could be reined in by carefully refining the mappings.²⁷

Precision. Focusing in more detail on specific aspects of control, participants discussed the system's precision and repeatability. The system was considered to have mixed levels of precision, in that the controller was very non-linear²⁸ and some combinations of control areas and mappings worked better than others.²⁹ Participant one felt that the ability for detailed control was its best quality.³⁰ Two people learnt to adjust to the levels of precision in the controller, both through adjusting their expectations³¹ and choosing mappings and sounds that suited the system's way of working.

Repeatability. Three participants considered the foam to have approximate repeatability,³² although very precise repeatability was difficult to attain. Repeatability was also affected by the transient and nonlinear nature of the mapping system, where the entire behaviour of the controller could easily change with small adjustments to the mappings. Participant one used visual feedback from parameter controls in Live to aid repeatability.³³

Intuitiveness. Three participants commented that they considered the system to be intuitive

²³Participant 2: 'I think the main thing with using it is learning to let go, because with all other controllers it's total accurate control more or less'

²⁴Participant 5: 'It was very unpredictable in a kind of a semi-annoying way'

²⁵Participant 5: 'It was more frustrating than anything, but sort of expressive but just frustrating in a sort of uncontrollable kind of way', participant 5: 'I think I'd just got frustrated with a it and I just thought it's not really doing what I want it to do'

²⁶Participant 6: 'in a weird way, the best thing about it is the semi randomness of it, and these kind of sounds that you can make with it, that I don't think you would make if you were trying each parameter on their own or that kind of thing, but then paradoxically at the same time it's kind of the worst thing about it because it means that you're not in control all the time', participant 2: 'I would say in a sense the unpredictability of it is a strength and a weakness'

²⁷Participant 1: 'You'd get to points which were really nice but then you'd also have a lot more things where there were problem bits, there's just more scope for things changing and getting suddenly loud or something like that. I think the more you use the more you have to put in time to refine it as well.'

²⁸Participant 2: 'A slight variation in pressure here or there can drastically alter the sound'

²⁹Participant 3: 'When you were exploring it and you found a certain area you could obviously lock into an effect by squeezing a certain part of the foam but you still can't rely on it like you can rely on where a key is on the keyboard. ... there's a general precision in terms of what you change depending on where you put it, but it's still not completely in control'

³⁰Participant 1: 'I think it was the best thing it worked for, just going into really detailed bits of the sounds', participant 1: 'I was at some points like squeezing with both hands ... and then just sort of gripping it and moving it very slowly once you've found those nice bits'

³¹Participant 3: 'it goes back to that thing about the nature of the device and the way that it is imprecise, and how much can you work with that? ... perhaps it's about understanding how much you can predict what it's going to do'

³²Participant 2: 'while it was on one sound, I got the parameters to where I wanted then messed around; I could, in that place for a bit, put it down let things settle and then I could kind of get there again, more or less', participant 4: '*if you found a hotspot could you return to that later?* If the EchoFoam software remains on the same setting, probably yeah. I wouldn't bet on it and but I probably would'

³³Participant 1: 'you could come back to it but what I found was that I always had to look at the visual feedback as well, and I was doing that to get to the same thing. So, not that that's a bad thing but just relying on how much I squeezed it before wasn't enough to let me get back to the same place, I had to sort of look as well and then moved around from there. But, yeah, I could definitely repeat things, in a sense not exact precision, but I could get to the same sort of types of things again'

in use; both in terms of the act of manipulating foam³⁴ and in terms of the way it could be used to explore timbre spaces.³⁵

Expressiveness. There were mixed opinions on how expressive the controller was, ranging from participants being undecided about it³⁶ to them feeling that it was very expressive.

³⁵ The degree of expressiveness was tied in with the range settings; carefully refined mappings could be used to limit or balance nonlinear behaviour, making the interface more playable and more musical.³⁷

Multiparametric Control. Multiparametric control was something the participants weren't accustomed to, having used more conventional MIDI controllers beforehand. Participant one commented on the advantages of being able to control multiple parameters in this style,³⁸ while participant two commented on how it was more difficult and took practice.³⁹ Three people talked of a balancing act to set up the parameters correctly for their patches, for otherwise moving one parameter could make another dependent parameter move to an undesired setting or move totally out of control.⁴⁰ This balancing act also made the system interesting to play, and led to the discovery of new and unintended settings.⁴¹

Bimanual Interaction. Everyone was asked whether they used the system with a single hand or both hands, to a range of responses. One participant would use two hands on the foam to explore a setting initially, and then move to single hand later. Others used two hand with the foam for most of the time. Three participants used the system as a whole bimanually, manipulating the controller with one hand while operating the computer with the mouse in the other hand.⁴²

6.5.4.3 SOFTWARE USE

Software Usability. The reports back from the participants on software usability for the 'EchoFoam Live Link' application were all positive, with only one minor bug reported.

³⁴Participant 8: 'It's the thing we know, we know the movement of it, we know how to squeeze something so that felt, that does feel very intuitive'

³⁵Participant 1: 'I like exploring in the sounds and it's not something that other controllers are that conducive to ... it just felt like it was a really intuitive way of doing it, and a very expressive way of doing that'

³⁶Participant 6: 'It was expressive in a way, in a very abstract type of way. I'm not sure to be honest'

³⁷Participant 1: 'it could be potentially very expressive interface but if used with a lot of thought and care I think', 'sometimes especially with smaller ranges set on the software you kind of restrict yourself anyway to those regions and then it kind of gets a bit expressive'

³⁸Participant 1: 'it was very layered and textual stuff that I was doing so it allowed me to be doing quite a lot at once. I could have just mapped it to a couple of things and then recorded that as a sort of single layer but I was doing a few things at once going on at the same time and then seeing how those could sort of intermingle and change over time'

³⁹Participant 2: 'Knowing what you're doing and what is affecting what, I think without a fair bit of practice to really follow mentally and I guess in sort of in a sense of physical movements, to follow that many sounds and movements all at once is quite a feat.'

⁴⁰Participant 8: 'Certainly, for some bits when I had things multi-mapped, there would be one thing that was where I wanted it to be, and then the other bit, when I would get to the bit in one sound that I wanted, the other bit could go totally out of control. So that was tricky but it was something to play against which was kind of interesting.'

⁴¹Participant 8: 'I would be trying to think about something else that it was mapped to at the same time, if I focus on that thing and then the other thing would go, so there's just a constant sort of balancing act. In a way, I guess trying to balance all of the parameters so it can make interesting sounds because there were always things that would come in against what you were consciously thinking about trying to get it to sound like. Yeah I like that.'

⁴²Participant 1: 'I was using it mainly in one hand I think ... While I was using it the other hand was doing click on or off things', participant 2: 'I was doing that with the mouse, tweaking them as I was squeezing to find the range I wanted.'

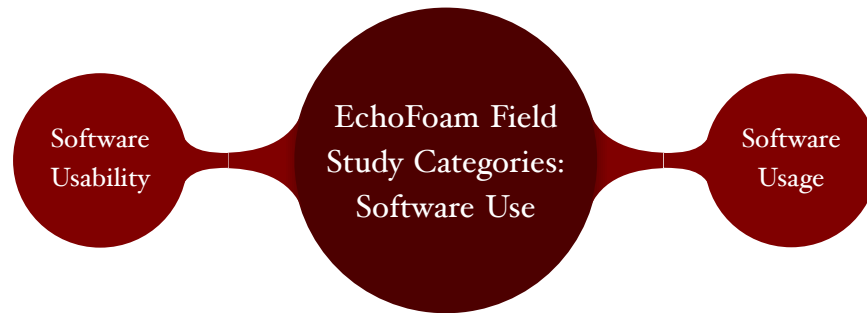


Figure 6.19: EchoFoam Field Evaluation Interview Analysis: Software Use

People commented that it worked well,⁴³ and was clear and intuitive in use.⁴⁴

Software Usage. Three participants discussed the particular features of the software they used. They used a range of features; some made use of everything including search functions and range editing while others used just the minimum mapping functions to connect the controller up to Live. The random population functions were popular for exploring settings quickly and providing starting points for further exploration.⁴⁵

6.5.4.4 MAPPING

Synth Mappings. Five of the participants felt that the system worked best for interacting with less precise,⁴⁶ and less rhythmic material,⁴⁷ such as slowly evolving ambiences and pad sounds.⁴⁸ There were a number of reasons for this; a pad can be a safe sound, such that less predictable control signals from the system will not necessarily stand out as sounding bad.⁴⁹ Pad sounds also match the physicality of the foam,⁵⁰ which is usually in motion as it expands back to its uncompressed position. Participant three felt that this relationship helped them to explore the controller.⁵¹ The speed of control with the foam and its less precise nature made it more difficult for controlling percussive sounds with

⁴³Participant 3: 'I think the software generally worked fine, no gripes at all.'

⁴⁴Participant 1: 'Yes it was absolutely fine and it was just intuitive, it worked well', 'The software is great ... I like the GUI.'

⁴⁵Participant 6: 'To begin with I just used the populate thing mainly for a while. Once I worked out which parameters were making the most interesting noises, like I thought when we were using it on the pads, when it was changing the transpose it was making some really nice artefacts, especially when it was going up really far, it was going past transposing it +24, so then I started to manually assign parameters.'

⁴⁶Participant 6: 'Some parts of it were really interesting, there was a whole soundscape that you could create. That's what I meant by that it was good for things that if you didn't really want to control them precisely and you just wanted things to happen ... For getting weird kind of abstract sounds it was quite nice'

⁴⁷Participant 4: 'however for soundscapes and stuff like this, longer sounds and then things with more like depth and less melody and rhythm, for those it was actually a lot nicer'

⁴⁸Participant 2: 'The main thing that I found that I quite liked to use it on was pad sounds, that's what it was ideal for, with sort of slowly evolving subtle textures, so for that I think it did work really well.'

⁴⁹Participant 3: 'you've got a consistency to a pad, and then if you use the controller with that, you're kind of altering things within that but you still have a general kind of sound going on underneath'

⁵⁰Participant 2: 'Going back to the start I mentioned it was always more appropriate for pad sounds. The sounds that you seem to create if you map it properly of course sort of mimic the motions that you are doing in the sense you know you've sort of, you can squeeze hard and it will do a lot smaller noticeable quicker things, slower movements all sorts of be reflected in the sound manipulation as well'

⁵¹Participant 3: 'using it with something like a pad enables you to morph a sound that's ongoing so you get more of an idea about what the controller's doing'

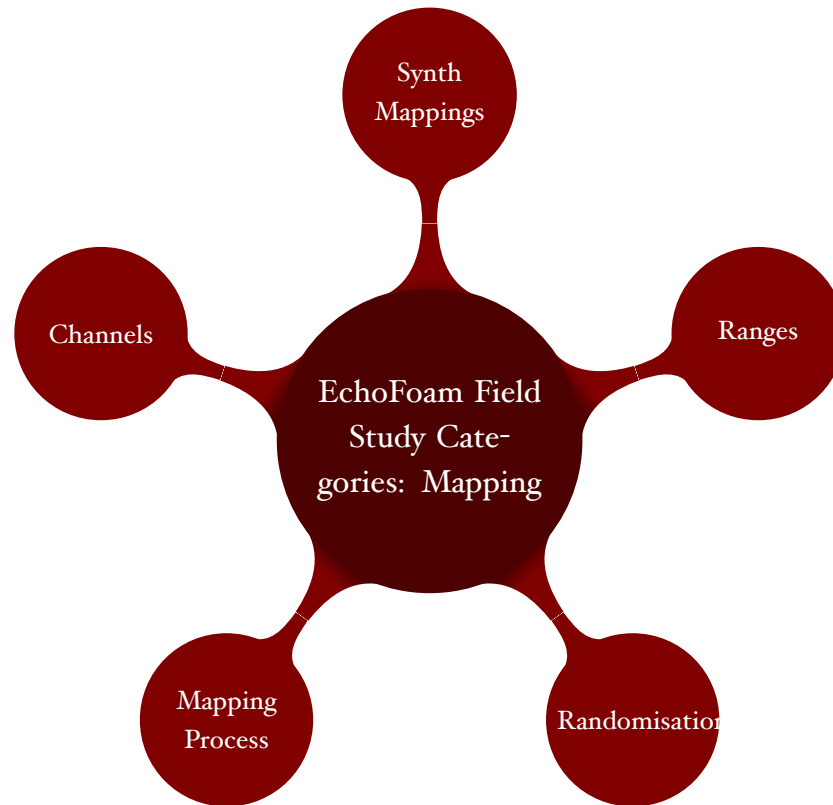


Figure 6.20: EchoFoam Field Evaluation Interview Analysis: Mapping

sharp attacks.⁵² Participant one preferred the control for more detailed mappings.⁵³ Participants gave examples of the mappings they used. These included:

- Mapping to analogue-style synthesiser parameters such as filter attack and resonance, envelopes and tuning.
- Mixing between wavetables.
- Controlling a variety of effects.
- Controlling spectral processing and buffer slicing in Max for Live.

Channels. Participants had the option of creating mappings for up to ten channels. They typically used the controller with a high or full number of mappings, although not all of these mapped channels were necessarily significantly affecting the sound.⁵⁴ Using a

⁵²Participant 2: 'I wouldn't even think to use it on percussive sounds or even affects processes on percussive sounds, partly because I want to have a lot of control, I want for the reverb on just that note I want it on, or send its two a delay channel, I want quite tight control. I would see it as sort of, it couldn't be discerning enough to do it in a musical way the, it would sound random sort of effect one. I really see it as being more in the sort of evolving textury[sic] tones more than anything.'

⁵³Participant 1: 'I think it was the best thing it worked for, just going into really detailed bits of the sounds'

⁵⁴Participant 6: 'Mainly I used it with the full amount of parameters and then used it on Ableton's 'Analogue' synth', participant 6: 'When you open it up it just comes with three automatically, so I started using that for a little bit and then it didn't take me too long to get up to eight, I played around a little bit and then once I'd worked out how to do it I just started playing around with eight. But then on occasion when there would be two or three that weren't doing much but I still liked the sound, I just left it. So it would probably be on the four or five that were actually affecting.', participant 1: 'I just used eight or ten, but not all of them were doing something interesting but I suppose maybe ... I used ten but maybe only like five or six were properly doing something'

higher number of channels could, dependent on the mapping, adversely affect control;⁵⁵ this was because the number of codependencies in the system increases exponentially as the number of channels increases, and the controller becomes more unpredictable. Participants tended to start using a smaller number of channels and build up to using more as they felt more comfortable with the system.

Mapping Process. There were several comments about how the mapping process needed to be approached with care to achieve good results.⁵⁶ This was because of the non-linear nature of control requiring more fine adjustment,⁵⁷ and the need to balance the codependent relationships between parameters.⁵⁸

Ranges. Following the importance of mapping refinement, setting good ranges was the cornerstone of this exploration. Participant four discussed how they would adjust each parameter manually to reach the optimum setting.⁵⁹ Correct ranges could make a seemingly unusable setting work, correcting for troublesome parameters such as envelopes or filters that could silence a patch if set incorrectly.⁶⁰

Randomisation. The use of randomisation in the EFL application proved to be a topic of interest in the study. Functions were available to create randomised parameter ranges, and to randomly populate the controller mappings with destinations from Live. Two people didn't use these functions at all; one didn't for functional reasons, because they were doing precise work. Participant two didn't use randomisation as a general philosophy, feeling that there was no pride in the results.⁶¹ Other participants did make use of the randomisation functions, two of them making it their typical mode of use.⁶² Four participants discussed how they enjoyed getting unexpected results; they also found it useful when they were stuck and needed inspiration, and enjoyed exploring timbre spaces that they wouldn't normally try.⁶³ Participant three described an example of this when the populate function mapped the controller to the course tuning parameters on a virtual analogue synth. They would never have normally manipulated these parameters when playing that synth and found the results compelling.⁶⁴ People talked of different workflows with the population functions; one approach was to continually re-populate

⁵⁵Participant 4: 'I had all of them and the lowest I think was four. I always went for quite a few actually, maybe that was also a mistake.'

⁵⁶Participant 2: 'the controller holds up its side of the bargain if you can get things rightly set up on the computer. Sometimes it wouldn't be as successful but I've found that some pretty much with a bit of ... as long as I got reasonable parameters with a bit of affecting the ranges then I could get it to ... represent my movements a bit better.'

⁵⁷Participant 1: 'it's something you really have to refine to get it working well but I think it's definitely something that's important specially with the nonlinear mapping'

⁵⁸Participant 1: 'I think the more you use the more you have to put in their time to refine it as well. ... if one thing gets to a nice place ... something else that was happening could get a bit out of control.'

⁵⁹Participant 4: 'just certain effects like phasors and flangers for example, it was really important to the change the range of each parameter'

⁶⁰Participant 2: 'I thought that the ranges were essential because lot of the time, just to keep it in a usable acceptable sounding point, because as soon as you would map something before you set any ranges, the control here would go 'bffffff' and that parameter would be no use, you might close a filter for example and you can't hear anything', participant 2: 'It seems the range settings are really important to stop this just waving about all over the place'

⁶¹Participant 9: 'Obviously there's a chance that your results could be great but there's no pride in having set something to random ... we as human authors can always appreciate efforts more than anything else.'

⁶²Participant 3: 'The populate function I found really useful in that it would apply random stuff and a lot of the time I wouldn't look to see what it was doing, i'd just keep repopulating it and squeezing it until it started making interesting sounds.'

⁶³Participant 4: 'The random thing was a nice idea but because if I had something that's had like an tons of parameters in it and like one of the VST's, and it would pick out all these ones that I wouldn't actually use.'

⁶⁴Participant 3: 'when you are using a pad and you use the octave shift and the frequency and the course tuning, because I would never use them. If I was composing and writing something normally I'd never think, I'd never

until a satisfactory setting was found. Another was to re-populate and then fine tune the setting, re-populating again if needed.

6.5.4.5 THE SYSTEM IN USE

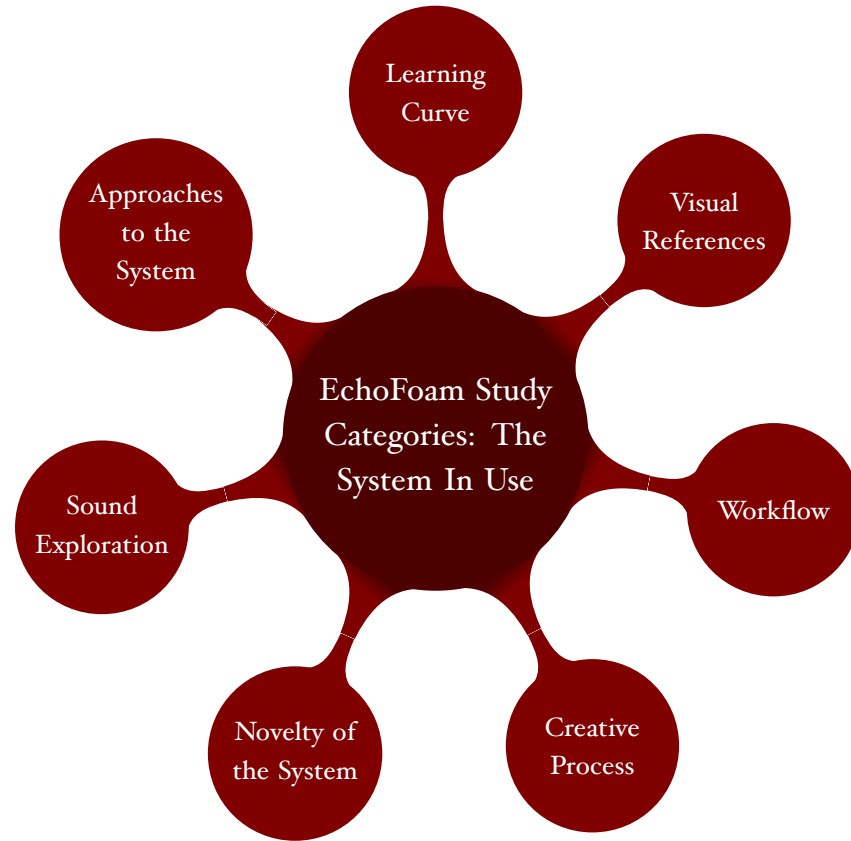


Figure 6.21: EchoFoam Field Evaluation Interview Analysis: The System In Use

Approaches To The System. Participants reported two main approaches to exploring the system. One was to treat it as a mechanism for discovering new sounds. The other was to explore its potential as an instrument.⁶⁵

Learning Curve. Feelings varied about the learning curve of the system. At one extreme, the controller was viewed to be quick to learn because the act of manipulating the foam felt intuitive.⁶⁶ At the other end of the spectrum, the foam was viewed as too inaccurate or random⁶⁷ and therefore difficult or impossible to learn properly. Frustration with the interface inhibited progression. Participant three felt that part of the learning involved

assign a controller like that where all those buttons were doing that thing at once. I'm not even sure that I ever write anything and use octave and pitch changes live, I'd always set them to something and then write it from that. Most of the time the things that I'd change would be filters, and envelopes so in a way that's been one of the main things □ when you are squeezing the pad and it get's that glitchy Ableton sound because it's struggling to do the realtime processing on the synth and then you get that real jumping up and down octaves, kind of extreme but it's something that I would have never done otherwise.'

⁶⁵Participant 2: 'In a compositional sense I didn't really use it as a sound finder, I was really just using it as an instrument unto itself'

⁶⁶Participant 1: 'It's the thing we know, we know the movement of it, we know how to squeeze something so that does feel very intuitive'

⁶⁷Participant 4: 'it's hard to define a learning curve when you can't really define at which point you would have mastered something random. It's the random factor which makes it hard for me to say that.'

being able to predict the accuracy of the device, adapting to the imprecision.⁶⁸ Three participants described a gradual progression in learning the system;⁶⁹ they started off with simple mappings and as they felt more confident or felt like discovering more they would make the mappings more complex to control. For participant four, the learning took place in using the software rather than the controller.⁷⁰

Workflow. Peoples' workflow patterns could be categorised broadly into two approaches: setting up then refining, and working in gradual increments. The use of the random population functions was popular for those taking the refining approach. A random setting would be generated and explored, after which any undesired settings could be changed to improve the sound.⁷¹ The incremental approach involved starting from a small number of mappings and targets, and gradually adding in more mappings to more destinations, refining these settings along the way.⁷² This incremental approach prevented the mappings from getting unmanageable or confusing by doing too much at once.⁷³ This process of adding settings was exploratory, and could lead to the discovery of new sounds and new tangents.⁷⁴

Visual References. While using the controller, three participants said they would mainly look at the parameters in Live. This gave another reference for the position of the foam aside from physical feedback, and aided with exploring the foam and with repeatability.⁷⁵

Creative Process. Participants were asked about how the system fitted into their creative process; the overall feeling was that the system was good as an exploration tool for discovering new sounds, and that it was very useful for exploring unintended paths, and for sparking creative ideas.

Four participants felt that its use as an exploratory tool was one of the most positive things about the controller, something that more conventional controllers are not con-

⁶⁸Participant 3: 'It goes back to that thing about the nature of the device and the way that it is imprecise, and how much can you work with that? On a level where you learn, perhaps its about understanding how much you can predict what it's going to do, learning in that sense.'

⁶⁹Participant 2: 'As each session goes, you start to get the feel of what movements are affecting what parameter in what way', participant 1: 'Even though it can be unpredictable and random, it's something that I think I would like to work with for a long time and develop'

⁷⁰Participant 4: 'the only path I could define as a learning curve was getting to know the software part. just understanding that. Then, after that... coming to soundscapes probably there was a bit of getting to know which ranges to use, being able to quickly look for where the major disturbance was when it wasn't sounding so right, I guess that took a little while.'

⁷¹Participant 6: 'The populate thing was quite good to get a grip of one pad or one synth, so it was really handy just to be able to go *boom* and that was that. You could sort of play with it a while and then, because sometimes one of the parameters wouldn't really necessarily be doing that much, so I'd manually change that one to the frequency of the first filter or whatever. That's what I mainly used it I think.'

⁷²Participant 1: 'I set up the live set and then I would just think about which things I wanted to map to it, map up a few, get those working, get the range going and then think OK, what else could I try linking it too as well, and then maybe choose some new parameters, take something off if it wasn't working so well, so just constantly kind of changing and seeing what things it worked for'

⁷³Participant 2: 'I did it in a sort of constructive sense; I would start off one at a time, the more that you add the easier it is to get lost, or maybe get one or two mapped in and then I would start searching for a range, be satisfied with that and then I would build, I wasn't really approaching it sort of add up and then muting channels off, I was just building extra, control by control until I was satisfied.'

⁷⁴Participant 2: 'if I've got something going and I may be made a couple of loops to go along with it, with my hands in a certain position I would stumble across the sounds I quite like'

⁷⁵Participant 5: 'my focus was on what sort of position my hand was in, because it's about emulating that sort of knob twiddling thing, it's about turning this foam square into a form that equals something, but it was about... my memory would be thinking about my hand is in this position and that knob is there but all those other knobs are in those places on Ableton as well, but I wasn't really looking at "I've got the foam like that", it was more "what's it doing to the screen"'

ducive to.⁷⁶ Participant six described how they would navigate sounds and pick out the best bits,⁷⁷ and participant one talked about how they enjoyed this process.⁷⁸

The systems' strongest function was for starting the creative process, the system lending itself to the discovery of unexpected results.⁷⁹ This would happen through the physical act of using the foam, and through mappings in the software. With high levels of complexity and a large range of possible outcomes, it was easy to come across new sounds.⁸⁰ The system in many ways forces you to do this in its natural workflow, where you set up mappings and explore them, and you don't know in the first instance what you are controlling.⁸¹ This unexpectedness is never eliminated from the system, still occurs at later stages when refining parameters. The populate function built on this to add more potential for creating ideas. Participant three compared the system to Schmidt and Eno's *Oblique Strategies*,⁸² a set of cards which give suggestions that may help an artist with new ideas or to overcome creative blocks (Eno and Schmidt, 1975, 1978). They emphasised the system's position in the creative process as a compositional aid.⁸³

Novelty of the System. Participants overall considered the system to be an unconventional approach to musical control, compared to conventional linear MIDI controllers. Participant two discussed this aspect of the system, commenting on how this approach was refreshing⁸⁴ and interesting.⁸⁵ This new approach could be difficult to adjust to as a new mindset was needed,⁸⁶ one participant commented that they needed to learn to 'let go' to use the controller.⁸⁷

⁷⁶Participant 1: 'it was completely new interaction ... because that's something I like exploring in the sounds and it's not something that other controllers are that conducive to □'

⁷⁷Participant 6: 'it's definitely a thing that you have to play around with for a good while and the pick out which bits were nice or which params were nice so its definitely an exploration type of deal.'

⁷⁸Participant 1: 'I like the kind of freedom it gave you to explore an area, you were in control of where that could go with the mappings'

⁷⁹Participant 1: 'it can make interesting sounds because there was always things that would come in against what you were consciously thinking about trying to get it to sound like. Yeah I like that.'

⁸⁰Participant 2: 'with a mouse you can only really control one thing at a time, it's a bit different if you have got something mapped to midi knobs, there are certain sounds that's you can only get when one thing it is set too low and one thing is set too high, you very rarely come across that, whereas this, just by chance and squeeze will throw that sound at you'

⁸¹Participant 6: 'you'd end up creating this soft of weird abstract sounds with it just because, too be honest, I didn't know exactly what I was controlling all the time which was kind of nice. So yeah I certainly thought that it has its purpose in this kind of semi-random fashion.'

⁸²Participant 3: 'The thought that I had of it mainly and the physicality of it and the way that it affects properties kind of randomly, it kind of reminded me of Brian Eno and his *Oblique strategies*. If you come to a dry patch in you composition, you go to them □ the foam reminded me of it in a way because it's the same kind of thing, especially with the populate function. If you were to get stuck or get bored of a sound or composition it's a quick way to find something new by accident'

⁸³Participant 3: 'I think it would be really useful as something to bring in to randomly change your sound for you and refresh it, help you find new sounds, ways of working that you wouldn't usually use, which is extremely useful because people search for that kind of thing all the time when they're composing, different ways to work and different ways to get stuff that they wouldn't otherwise come across. I know for me personally, composing, I can kind of get very stuck in a certain way, it's really easy to get a system of composing that works well enough for you not to feel that you have to change it, but then sometimes I think you really need something that will just put you out of your comfort zone, and find new ways of working that you wouldn't otherwise do. I think that's definitely the best thing about it for me'

⁸⁴Participant 2: 'The approach to using the controller is important. I've been approaching it with the o-127 midi knobs and on off button slider mindset. Those are clearly better controllers for this type of control. Why bother re-creating it. It actually a refreshing approach to think a bit at the start to set up acceptable parameter and usable ranges. Then just give it a squeeze.'

⁸⁵Participant 2: 'it's novel and new and different, like a totally different approach, I like it, for me that's quite an interesting thing.'

⁸⁶Participant 2: 'Overall it's a totally different approach so it takes time to figure out exactly what you're doing, it's fun'

⁸⁷Participant 2: 'I think the main thing with using it is learning to let go, because with all other controllers its

6.5.5 DISCUSSION

6.5.5.1 METHODOLOGY

Before analysing the results from this study, a discussion of the methodology and its effectiveness will help to put them in context. Overall, some very rich and detailed results were collected from the six participants who completed the study, leading to a large set of interview transcriptions for analysis. The participants had a variety of aims for using the system, and used it in a variety of musical scenarios. These findings could not have been discovered in a controlled, laboratory based study. An example of this is where one participant discussed using the controller in an improvisation session with friends, the insights this gave them into the system and how this changed their opinion of the controller. Only a study like this can encompass this kind of spontaneous use.

Some aspects of the study did not go as well as possible. Unfortunately 40% of the participants did not complete the study, due to varying circumstances outside the control of the researcher. While obtaining results from all ten participants would have been ideal, obtaining 60% has still given sufficient data about the controller. Further to this, because of the very open ended nature of the study, there are some areas of use and issues raised where the participants followed different paths and did not cross over. The results describe the stronger, saturated threads in the results, but because of this reduced overlap, there are some weaker threads which, while compelling, were not significant enough to consider in this analysis. They would be interesting to follow up, though that would be outside of the scope of this thesis. One answer would have been to give the participants a more restricted piece of software to use for the study, but this would have had less chance of capturing their interest over the course of the study. A larger number of participants may have helped with this issue of overlap (if the time and resources to support this had also been available), but may also have increased the number of weaker threads in the results. Lastly on the subject of participants, all subjects were students, and involving other non-academic musicians in the study would have expanded the scope of the results.

Finally, participants did not submit as many supplementary materials as was hoped for. They were asked to keep a record of their experience with diaries, sound recordings, screen shots and so on. While two participants did submit some data (see section 6.5.2.2), others failed to keep records, and this highlights an issue in the design of the study. The collection of these additional materials relied purely on the will of the participants, who were already giving up a significant amount of time for little material reward. Some sort of motivation scheme may have been put in place to encourage record keeping, but this may have conflicted with the ethos of the study which was not to interfere with people's creative process and environment. In future studies, any motivation measures would have to be carefully chosen to preserve this balance as much as possible. As there was little supplementary data recorded, the majority of the results are from the participants' recollection of their experience in the final interview.

To summarise, there have been methodological issues which have detracted from the completeness of the results. Overall however, the study has yielded a large set of useful data, which

total accurate control more or less, this is not that approach so that was probably the biggest challenge, just a totally different mindset'

has helped to build a compelling picture of the EchoFoam system *in the wild*.

6.5.5.2 ANALYSING THE RESULTS

The controller provoked a mixed range of reactions from the participants. This went from fairly negative from two participants to a much more positive view in the rest. Certain elements were more enthusiastically received than others; the differences in reactions to these elements reveal some interesting general points about the participants' use of digital music controllers and how they fit into their creative processes.

From a functional point of view, the system as a whole proved to be very robust. None of the controllers malfunctioned during the study, and the reactions to the software usability were all very positive with only very minor complaints. From this, it follows that the results are reactions to the system itself, rather than being due to usability problems.

A very prominent theme was controllability. Participants found varying degrees of instability in the system, originating from three possible sources. Firstly, the foam itself; the physical material is always trying to return to its original form and therefore always in motion for a period after you manipulate it, which creates possibly unwanted control signals. The foam is also complex to manipulate, requiring the user to work in three dimensions rather than typical one dimensional controls used in scientific-style music controllers. This makes it difficult to achieve more precise results without practice. Secondly, the multiparametric and nonlinear nature of the controller means that there are always surprises, and forces the player to adopt an exploratory strategy for each new mapping. Interestingly, some participant perceived this nonlinearity as randomness, although the mappings were in fact consistent from initial random settings. Thirdly, the system would become more unstable as more channels were added, due to increased complexity in the mappings and increased possibilities in sonic output. The participants generally found that the unstable elements of the system could be brought into line with carefully refined mappings, and this seemed a successful overall strategy for using the controller. This issue of controllability underpins all other prominent themes in the results.

Given the unconventional nature of the system, some participants found it difficult to make the shift in perspective towards the exploratory approach needed, or simply disliked this mode; a couple of participants had a strong preference for linear, separable controllers. Participants who achieved the most with the controller took an emergent approach, exploring what the controller could do for them rather than trying to impose a preconceived function onto it. This variation in participants' approaches could be due to the cognitive styles of individual participants (Guilford, 1980). This theme is explored further in section 8.4.3 of the conclusions.

A strong thread in the results concerned the system's place in the creative process, which was widely considered to be at the idea generation phase of creating a piece of music. Participants found the device to be best for finding unknown sounds, and for creative inspiration. The system was indeed designed for this kind of work in mind so this is not a surprising result, but it is interesting to analyse why this is the case. The key feature is that there is a degree of unpredictability about the system, in the way the mappings are set up and the way that

the physical controller works. Most controllers and sound engine combinations have some degree of unpredictability in the output, but in this system it is made explicit. The output of the controller is complex so you must explore new settings to find out how the motions map to the sonic results, to discover how each motion changes the sound. Further to this, the system's precision forces it away from the more subtle refinement stages of composition. Only one participant felt that it was precise enough to achieve very subtle results, the others feeling that it lacked enough accuracy for detailed work. The use of randomness in the mapping process strengthened the system's place at the idea generation phase of the composition process, and proved to be popular with some participants who enjoyed exploring sequences of random sounds and refining the results.

Something that all participants agreed on was that the controller was at its best when mapped to the control of slowly evolving textures and ambiances. This reflected the physicality of the foam. Further to this, these kind of sounds provide a constant reference when exploring unpredictable mapping landscapes.

6.5.6 CONCLUSIONS

Having analysed the data from the field study, the following conclusions can be made:

- While some people enjoy multiparametric, non-linear control, others dislike it, preferring linear, separable controllers.

This reflects and adds weight to the same conclusion in the timbre space exploration study (see section 7.8).

- Imprecision and unpredictability were both hindrances and the greatest strengths in the system.

These factors led to unintended use and creative inspiration. They also limited the system's use as a refinement tool.

- For successful use, a considered approach to mapping was required.

Without careful mapping, the system could become unmanageable in some cases. Unpredictability and imprecision could be compensated for with careful attention to parameters, leading to use for more detailed work.

- The system is most useful in the idea generation phase of the creative process.

Participants enjoyed using the EchoFoam controller for creative inspiration, and for finding sounds that they wouldn't normally come across.

- The system has its place among other more conventional controllers in the computer musician's toolset.

It provides a type of control that is not possible with conventional MIDI controllers, and is a useful tool to complement them.

6.6 Summary

This chapter has described the development and evaluation of EchoFoam, a system for malleable, musical control. The evaluation results, combined with the results from the Phalanger

studies, contribute towards a wider picture of multiparametric control. These results will be considered together in chapter 8. Before this, there is one final study to present. This is a formative study of the RECZ system, a meta-mapping system for multiparametric controllers, designed for the exploration of large parameter spaces. This study used both Phalanger and EchoFoam controllers; it reveals some points of comparison between the two controllers, and provides further data for the discussion of multiparametric control.

CHAPTER 7

MULTIPARAMETRIC EXPLORATION OF TIMBRE SPACES

A lot of our experience with gear is, if you know it well enough, no matter how simple it is, you can find something that it perhaps wasn't intended for.'

Rob Brown (Richardson, 2008)

7.1 Introduction

This chapter explores the application of the Phalanger and EchoFoam controllers to a specific computer music problem: the exploration of timbre spaces. In this context, timbre space exploration is the act of navigating through the different possible outcomes of a sound synthesis process, parameterised by multiple, continuous controllers. Such a parameter space may have several dimensions, resulting in a huge number of sonic possibilities, and a tool to help navigate this space intuitively can be extremely valuable.

This exploration task was inspired by the first EchoFoam evaluation study, where participants haptically explored six dimensional timbre spaces. The results suggested it would be interesting to test high-level strategies for controlling the subset of parameter space within which the controller was working, in order to help the user explore the full space more effectively. The results from the study also showed some participants felt the behaviour of the controller to be unpredictable and imprecise in some cases, so imposing a larger controllable constraint onto its output could complement the behaviour well, making an interesting combination of order and unpredictability.

Both controllers are very similar on a data level in that they stream continuous, multiparametric, codependent data. This means that they can both be plugged into the same software process, creating an interesting opportunity for comparison. The comparison would be especially interesting considering the two opposing styles of tangible and intangible interaction.

The continuous outputs were mapped to continuous synthesis parameters to allow exploration of the spaces. With Phalanger, moving or changing the shape of the hand changes the sound. Using EchoFoam, deforming the shape of the controller changes the sound. The software mapped its inputs through an ESN in order to change its dimensionality to match the number of synthesis parameters being controlled. The controllers become too sensitive with larger parameter spaces, so a navigation system was developed to enable high level control over the subset of the parameter space in which the controllers are working. By moving and refining the working range, a timbre space can be progressively explored to find a desired sound. The search process was developed by focusing on three progressively more complex scenarios. The first tested a simple scenario, using a four dimensional parameter space. The second scenario tested a higher level of complexity using ten dimensions, and the final scenario explored if and how one might control a huge space with forty dimensions. The system is used bi-manually, while one hand is used for detailed search with one of the input devices, the other hand controls high level search parameters with MIDI and the computer keyboard.

The system has been named *RECZ*, after the four principal operations involved in its use: randomise, explore, centre, and zoom.

Reactions from two musicians aided the initial development, after which an evaluation study was carried out.

7.1.1 RESEARCH QUESTIONS

The project was approached with the following questions in mind:

1. Realtime control of the size and position of a subset of a set of parameters: is this a useful way to navigate a timbre space?
2. How does this approach compare to using conventional editing tools such as a GUI or knobs and sliders?
3. How do the two controllers vary in their ability to control this process?
4. Is it possible to interactively explore a continually controllable synthesis algorithm without knowledge of its underlying workings?

7.1.2 UNPREDICTABLE CONTROL

A motivation of this study was to explore accuracy and unpredictability of control. Gelineck and Serafin (2009b) commented on this issue in their survey of electronic musicians. 16 of 18 of the musicians they interviewed said they preferred tools that they don't fully understand or are unpredictable in some way. The two interfaces used in this project have accuracy proportional to the skill of the user and were perceived in previous studies by some users as unpredictable and imprecise compared to conventional controllers. It is compelling to see how this style of interaction can fit into the more precise world of digital music by imposing a structured framework. By applying these two controllers to the task of timbre space navigation, this issue is explored further.

7.2 Related Work

In the area of timbre space navigation, evolutionary methods (e.g. Dahlstedt, 2007) have been well researched. Using interactive evolution, synthesis patches are encoded as individuals in a large population and evolved with genetic techniques to navigate the search space. Although effective, these techniques suffer from the bottleneck of human evaluation of fitness, which reduces the efficiency of the process. Seago, Holland, and Mulholland (2008) explored timbre space navigation, proposing two strategies; multidimensional line search and the use of adapted Bayesian filters. A controller-based approach is taken by Van Nort and Wanderley (2007), who used a graphics tablet and mapping engine to navigate complex sonic spaces.

This project fits between these areas, applying multiparametric controllers to the problem of timbre space navigation. A study by Dahlstedt lies in a similar area to this work (Dahlstedt, 2009). He investigates multiparametric mapping strategies, controlled with a variety of MIDI devices. He tests two strategies, one based on navigation through the parameter space with vectors, and the other based on a gravity analogy. Like the techniques used in this study, both allow meta-control of the scaling of mappings across the timbre space, and both use randomness for initial settings. Dahlstedt names this technique *dynamic mapping*.

The initial development of RECZ will now be described.

7.3 System Architecture

The timbre navigation system consisted of a collection of different programs and modules. The programs that run with each controller exist as separate OpenFrameworks applications, so an overlay application was designed that could host the timbre search engine and be integrated with both pieces of software. Figure 7.1 shows an overview of how the system was setup. Ableton Live was chosen as the sound engine; the software communicated with Live using liveOSC¹, an OSC/Python interface that runs as part of Live. This provides a convenient way for the timbre navigation system to query the amount, type and ranges of synthesis parameters in Live, and automatically assign control streams to them.

7.4 Exploring Timbre Navigation Strategies

What we have already is the input from two different controllers, both of which can control an arbitrary number of parameters in an arbitrary way. With the foam, the user explores the parameter space by deforming its shape, and with the hand tracker the user can manipulate parameters by moving and changing the shape of their hand. In both cases, the sound will change in direct relation to body movement. When controlling larger subsets of the parameter space, small motions amount to very large timbre changes, so in terms of controllability, working in a smaller subset is better for fine tuning of the sound. This means some kind of strategy for moving and narrowing the subset is required. Both controllers also output code-dependent parameters, which means that with direct mapping it will not always be possible to reach every available timbre, so a strategy is required to compensate for this.

A timbre space navigation engine was designed with these issues in mind, the development of which is now described. The development took place over three progressively more complex scenarios. Overall, a *blackbox* approach was used; the system was designed with the aim of navigating any synthesis process with continuously controllable parameters, purely with controllers and no GUI.

7.4.1 SCENARIO I: A FOUR DIMENSIONAL TIMBRE SPACE

The control streams were mapped to four parameters of a virtual analog synthesiser. For this scenario, the core of the navigation process was designed. To create variable subsets within which to explore with the controllers, each controlled parameter was assigned upper and lower bounds to create a working range, and also a polarity to determine if the mapping would be inverted. Pressing a key on the keyboard randomised these ranges and polarities, creating random areas of the timbre space to explore. The next step was to facilitate control of the subsets, to allow gradually refined navigation through the space. To achieve this, first a global percentage multiplier was applied to each range, so that the ranges could be narrowed around their centre points. A continuous MIDI controller was mapped to this. Secondly, a mechanism was introduced to move the centres of the ranges; pressing a key on the keyboard caused

¹<http://livecontrol.q3f.org/ableton-liveapi/liveosc/>

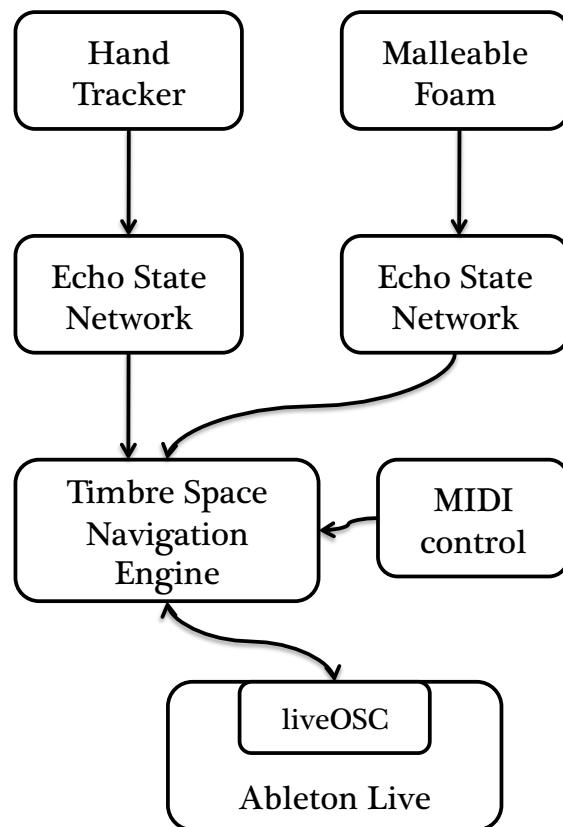


Figure 7.1: System Architecture

the ranges to centre on the current value of each parameter, allowing navigation through the space. With these functions, the workflow for navigating the parameter space was as follows:

1. With the ranges set at full, explore different random subsets of the timbre space until something in the area of the desired result is found
2. Centre on the area of the desired sound and zoom in a little by shrinking the working ranges
3. Explore the new ranges, getting closer again to the desired sound
4. Repeat the last two steps until the final result is reached

One issue was that using uniform randomness to determine the ranges sometimes led to tiny ranges on one parameter, limiting the scope of exploration. To remedy this, the random number generator was mapped through a sigmoid curve, making larger ranges more probable.

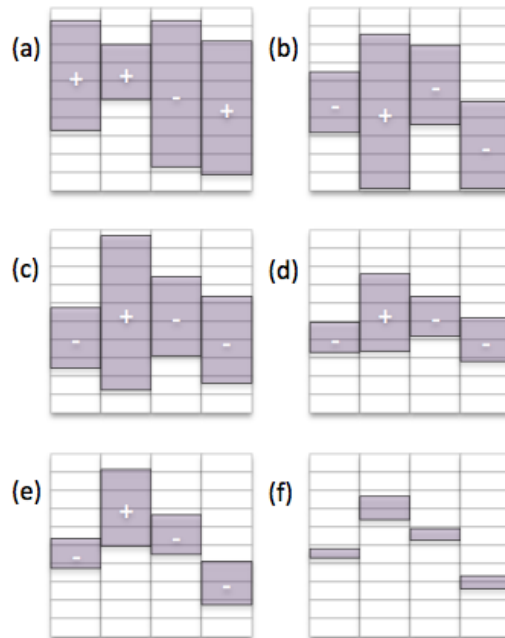


Figure 7.2: Exploring A Timbre Space

Figure 7.2 shows an example of how a four dimensional timbre space might be navigated. In each of the six steps, the columns represent four synthesis parameters, the shaded boxes represent the working range for these parameters and the + or - indicates mapping polarity. In (a) and (b), different random ranges are trialled to reach a suitable setting to start refining from. In (c) the ranges are re-centred and in (d) they are scaled down. In (e) and (f), centering and scaling is repeated, arriving at a small subset of the possible space.

This solution showed initial promise for navigating this smaller timbre space. The next challenge was to try exploring more dimensions.

7.4.2 SCENARIO 2: A TEN DIMENSIONAL TIMBRE SPACE

The system was mapped to ten parameters across two plugins, a sampler and a chorus that processed the sampler's output. With this number of parameters, the process still seemed to work effectively. Up to this point, each output from the controller was mapped to the same parameter on the synthesiser. However, given the nonlinear nature of the controllers, each output stream behaves differently from another, which may limit the extent of the search space which can be navigated. To solve this, when the ranges were randomised, the parameter targets for each output stream were now randomised as well. To enable further control for the navigation process, code was added to enable mutation; pressing a key on the keyboard caused Gaussian randomness between -10% and 10% to be added to the bounds of the ranges, allowing subtle variations in the exploration space. The addition of these new features helped further improve the control over navigation. The next scenario to explore was the case where a synth had more parameters than could reasonably be output from the controllers.

7.4.3 SCENARIO 3: A FORTY DIMENSIONAL TIMBRE SPACE, NAVIGATED WITH 10 CONTROL STREAMS

The VOPM² softsynth was chosen for this scenario, an FM synthesiser with over forty parameters; forty of these were selected for control by the navigation engine. This was an interesting challenge as it involved nonlinear control of a highly nonlinear soundspace, and also because FM suffers from difficulty in mapping between gestural input and synthesis parameters (Lazarini, Timoney, and Lysaght, 2008). At first this scenario seemed to produce no sound on many settings. It was found that this silence was caused by one key parameter when the value was above 20%, so this parameter was removed from the set of targets.

As there were fewer control streams than parameters to control, the target parameters were selected as a random subset of the available targets, and could be changed again randomly by pressing a key on the keyboard. Selecting a new random set of targets left the previous targets on the values they were at when the settings changed, so each new random jump navigated further through the timbre space. The nonlinear nature of this timbre space meant that some settings were silent or would jump very suddenly to a different sound, so a one level undo function was added enabling the user to jump back from an unwanted setting. Navigating through a larger set of target parameters with random target selection in this manner allowed each new setting to be explored in an embodied way with the controllers, and rejected with the undo function if the new set of targets didn't take the user in the right direction. To widen the search options further, a function was added to randomise the ranges of the currently selected targets while preserving the unselected ones.

7.5 Initial Reactions

Some initial thoughts about the system were gathered from two musicians, during informal interview sessions. During both interviews, debugging information was showing on the screen

²http://www.geocities.jp/sam_kb/VOPM/

at the start; both musicians found the experience to be much improved with the visuals removed so they could concentrate on the controller. Most importantly, both musicians found the system engaging to use. Both preferred the hand tracker as the controller, finding it easier to keep points of reference. One attempted to navigate from a distorted sound to a clean sound and back again, and achieved this successfully. They found that with the hand tracker they could return consistently to a previous point in the timbre space, and felt that their ability to do this would improve with practice.

One issue was with using the range centering process; when the working ranges are moved to a new centre, the position on the controller now corresponds to a new position in the parameter space so the sound changes. This can disrupt the flow of navigating the search space, as it's sometimes difficult to find where the sound you had centered on is in the new search space, although it should always be possible to find it. Another issue was with the size of ranges, one interviewee felt it would be better to always start from much wider ranges or a completely full range.

7.5.1 DISCUSSION

The initial reactions demonstrated that the system could work successfully, although some refinement was needed. The main issue was the interaction between the range centering process and the current state of the controller. Strategies need to be found to smooth out this process which in turn will improve the flow of the navigation experience. Another issue was widening the random range selection which could be solved by providing MIDI control of the sigmoid curve which maps the random range values; the user could determine how likely ranges were to be large.

One interviewee commented that the system was good for making broader adjustments to the sound, but really fine adjustment was difficult; bearing this in mind it's interesting to consider where this type of system fits into the composition and editing process. Gelineck and Serafin (2009b) observe that musicians need more accurate control when they come to the final stages of a composition, so this system may fit in best at the earlier stages of creative exploration. Any settings discovered with the system can be fine tuned with a mouse and GUI later.

In terms of control, an interesting property of the system is the use of randomness; random values are used to move around the search space in search of a good place to begin fine-tuning parameters. This is necessitated by the blackbox approach to parameter control and also by the nature of the controllers. To determine mappings by something other than randomness, for example manual control, would require the attachment of meaning to variables in the system in relation to the sound being controlled, however given the embodied nature of the controllers, meaning in this system is derived from listening and physical interaction in an explorative process. Considering this, using randomness seems to be the most appropriate approach, although varying the distribution of randomness, for example with sigmoid or Gaussian mapping, can increase the level of control.

An interesting aspect of this system is the role of bi-manual hand use. Treadaway (2009), discussing hand use in creative practice, describes how in manual activities the dominant hand

is used for micrometric and internally driven actions while the non-dominant hand is used for macrometric and externally driven actions, reflecting differences between the left and right brain hemispheres. This pattern of hand use is echoed with this system, one hand being employed for detailed exploration of the sound space while the other controls meta-level search parameters.

Returning to the questions posed in the introduction, on the issue of whether this system can plug in to any continually controllable synthesis process, a key issue is the selection of parameters. In any synthesis engine, as observed in the FM synthesis scenario, there are certain parameters (for example master volume) that hold significance over others and should be excluded in a timbre space search. Parameter selection also depends on the intended sequencing of the sound. For example, including envelope attack in the search space for a sound played as staccato would not be relevant. Setting the initial target parameters is something that could be controlled by a GUI, and is part of the wider creative search process.

Two questions in the introduction concerned the comparison of the two controllers used with the system and also the comparison of the system with conventional control methods; there is not enough data to answer these yet, and a formal user evaluation will help to find some answers.

7.6 Evaluation

Having received positive feedback from the informal feedback sessions, it was decided to carry out a more formal laboratory based evaluation, focusing on core aspects of the system. This study took place in spring 2010.

7.6.1 PARTICIPANTS

Six people took part in the study, (two female, four male). Participants were advertised for on university mailing lists and through word of mouth. Four of the participants were music students, one was a DJ who used digital tools, and the other was a researcher who used digital music tools. Participants had received between zero and nineteen years of musical training, with an average of 7.2. All were experienced users of at least one computer music application. Two of the participants had taken part in the previous evaluation of the EchoFoam controller.

7.6.2 SCENARIO

There was a significant amount of information for the participants to take in during the interview session; in learning how to use each controller, and in learning how to use RECZ. To try and simplify the study as much as possible, a single scenario was set up for the participants to explore. They tried this with each controller at two different levels of complexity.

Each controller was mapped through a trained ESN to output ten streams of data. These streams were mapped via MIDI to a patch in Ableton Live. The patch consisted of a sampler processed by a chorus, playing a major-7 chord pad sound. The chord played continuously, while the participants could manipulate sound parameters with the controllers.

The ten controller destinations were:

1. Sampler filter frequency
2. Sampler filter resonance
3. Sampler filter morph
4. Sampler shaper amount
5. Sampler FM modulation amount
6. Chorus delay time
7. Chorus LFO amount
8. Chorus LFO rate
9. Chorus feedback
10. Chorus dry/wet

This combination of parameters gave an expansive timbre space to explore, with plenty of variation between possible sounds.

The high level search parameters were controlled with a laptop keyboard. A single slider on a Kenton ControlFreak MIDI controller was used to scale the working range.

7.6.3 METHOD

There were two phases in each session. To begin with, the participants were asked to explore each controller, without using the meta-search functions. They were asked to report when they had a feel of how the controller worked and were ready to explore more functionality, after which RECZ was explained to them. They then experimented with each controller together with the RECZ. After this, they were interviewed about their experience. To allow for individual differences between participants, there were no fixed timings for each phase of the experiment. The sessions lasted for between 30 and 45 minutes. Interviews were semi-structured, and recorded as audio, lasting for an average of fifteen minutes. The transcriptions were analysed using a grounded theory approach (see section 4.3.5.2).

7.6.4 RESULTS

The interviews were coded according to grounded theory technique, and the resulting concepts fell into five categories: the system, the search process, control, player experience and methodology (see figure 7.3). The results will now be described, along with some supporting quotes where they supply additional detail.

7.6.4.1 ABOUT THE SYSTEM

Two participants who had taken part in the previous evaluation of the EchoFoam system (see section 6.3) commented on how this compared to their previous experience. Participant six

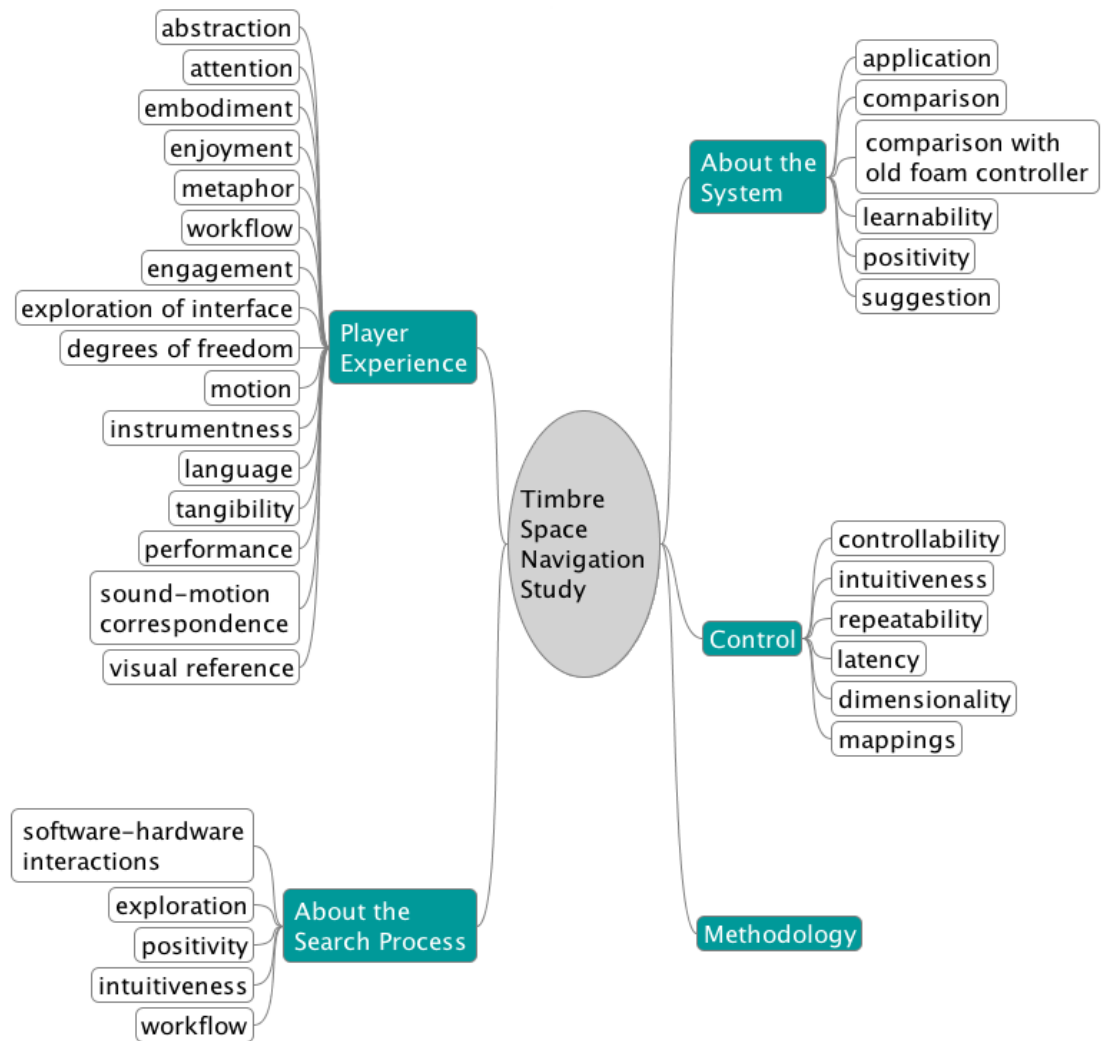


Figure 7.3: Results: Concepts and Categories

found the controller still to be familiar from the previous session.³ They also felt that RECZ had enhanced the experience of using the EchoFoam compared to when they used it first.⁴

All participants made positive comments about the system. For example, participant two enjoyed the playful nature of interaction,⁵ and participant five enjoyed using their hands in this way.⁶ Participants two and three considered the controllers to be easy to pick up⁷ although participant four had trouble adjusting to the foam.⁸

Some interesting comparisons were made between the two input devices, and in general the qualities of Phalanger were preferred to those of the EchoFoam controller. Comments were made about Phalanger being better in terms of latency,⁹ freedom of motion,¹⁰ ease of interaction,¹¹ repeatability¹² and immersion. One participant felt the foam was better for fine control when used in conjunction with the search process.¹³

Several application contexts were suggested for the system. These included interacting with live instrumentalists, recording human motion with EchoFoam, control of guitar signal processing, and the control of DJ effects. Participants made some suggestions for improvements and additional features. One suggestion was for the use of two hands for controlling Phalanger, in a sculpting metaphor. For RECZ, parameter freezing was suggested, so that individual streams could be frozen as the search came closer to the desired sound. Navigation way points were another request, allowing the user to bookmark points in the parameter space to return to later.¹⁴

7.6.4.2 CONTROL

In terms of general controllability, a level of subtlety and detail could be achieved with the system.¹⁵ In particular with Phalanger in conjunction with RECZ, one participant described how they made broad and then detailed changes to the sound.¹⁶ People outlined some of

³Participant 6: ‘... there’s a familiarity you can come back to. Even though it’s thicker and newer, this foam interface, I still know I didn’t have a lot of adjustments to make because I knew the interface even though it was different’

⁴Participant 6: ‘If I had to say in terms of hierarchy, that’s much better having those extra dimensions. In fact, I think that enhanced the exploration level drastically, I would say maybe more than 200% because it doesn’t end, the exploration doesn’t. It’s an infinite space in a way.’

⁵Participant 2: ‘I like it. I like how it feels. I like the idea that i’m playing with nothing and yet something is happening.’

⁶Participant 5: ‘especially with the hands it worked really really well’

⁷Participant 2: ‘I think a non-musician might really get on with that’, participant 3: ‘The thing I liked was... I felt I needed less knowledge beforehand’

⁸Participant 4: ‘I didn’t really know which bits were doing what so to begin with, there wasn’t much feedback in terms of where you’re putting something’

⁹Participant 3: ‘I like the camera a lot more. I felt I had a lot more immediate effect.’

¹⁰Participant 6: ‘With the foam I found that it was mostly pressing on areas or turning it in certain ways so I found that this was more constrained’

¹¹Participant 5: ‘For just playful playing around I kind of preferred the hand to the foam. It made it easier, more immediate and more embodied’

¹²Participant 5: ‘I know it’s not a direct mapping but never the less I thought it had more regions that I could go back to and do gestures that I could reproduce more easily than with the foam’

¹³Participant 4: ‘Without the additional faders I preferred the camera, but I felt when I was using the faders, the foam had a little bit more fine control’

¹⁴Participant 5: ‘Maybe another addition to that would be to be able to store some of these centres so that you could come back to... you could jump around these presets and then open up the design spaces around them’

¹⁵Participant 4: ‘you start to lock on a sound and really subtly move around that point’

¹⁶Participant 6: ‘And then I was thinking in the end I was trying to start making a little sound track. The dog barks, the wind blows... when I found the voices of gaza I thought that’s quite dark so lets find like Arabian women

the initial difficulties in controlling the foam, including repeatability¹⁷ and holding the same sound. The inherent latencies in the foam were a problem for three participants,¹⁸ but participant six felt that they adjusted well to this after an initial period; they felt comfortable with both interfaces after exploring their constraints.¹⁹ Phalanger was generally considered to be easier in terms of repeatability, however, the search process was considered to enhance the repeatability of the foam by one participant who had tried it before.²⁰

There were several comments about the multidimensional approach that the system takes. Participant 6 considered it to be intuitive in use,²¹ and two participants valued the holistic approach to exploring sound.²² Participant four considered the system to be expressive, although participant three found the process difficult to control²³ and participant two felt there should have been fewer parameters.²⁴

7.6.4.3 ABOUT RECZ

Reactions to RECZ were generally very positive.²⁵ The system was considered as a good complement to the two controllers,²⁶ and was found especially useful to compensate for control issues with the EchoFoam.²⁷

Specific aspects of RECZ were discussed. Starting with the randomisation or search spaces, while participant two disliked randomness in general,²⁸ others found this was a useful way of exploring different aspects of the timbre space.²⁹ In particular it was a good method

singing, so then I was slightly shifting them, shifting in that soundspace to make them more friendly'

¹⁷Participant 5: 'When I was just playing around with it I had very little means of just understanding the ranges of the different parameters and arriving at similar sound again'

¹⁸Participant 5: 'I thought that it's slightly annoying that there is that lag in the foam and I know that there's nothing that you can do, it's just built in because it needs to expand and contract.'

¹⁹Participant 6: 'I think it takes a moment to understand how fast you can go with it. It takes a moment I think for both interfaces to just get used to each other and see what kind of pressure you can have, what kind of speed you can have but then once you're in it there's a familiarity you can come back to'

²⁰Participant 6: 'Yeah I think this really works in terms of getting repeatable sounds. I don't know what I said in my last interview, but I do think that I said something about being not able to so easily getting to the same sound again, whereas this time around I think it was much easier. And I think it was to do with the dimensionality and it has to do with being able to be more in control of finding this place you like and exploring it'

²¹Participant 6: 'I think it's very intuitive, I would prefer that to any knob twisting or twirling. In fact I think both of those interfaces are extremely intuitive and the dimensionality is... you don't really think about it, it becomes something you work with'

²²Participant 5: 'To a certain degree, at least for me, I sort of stopped thinking about the single parameters, I became more a place, and I think that's much better anyway if you have that number of parameters, if it is just one space that you navigate, you stop thinking about different parameters or dimensions in that'

²³Participant 3: 'it was difficult to maintain one part of a sound and change other bits'

²⁴Participant 2: 'I feel perhaps its current implementation, maybe there are too many parameters being accessed that my feeble mind can't really grasp and find it a little bit too random in nature'

²⁵Participant 2: 'I think really that addition of that search function did a lot to the way, how engaging and how enjoyable it is to play', participant 6: 'How did the search process feel in general? Very intuitive. It was gratifying.'

²⁶Participant 5: 'the search function really adds to the way you play this because... with a cello, if you put a finger somewhere you kind of know what to expect but with a very complex multi-parameter synth you kind of change your hand slightly... you're never quite sure whether you can reproduce what you've just done so with that search function its really nice'

²⁷Participant 5: 'the feature that you can zoom right in kind of takes you back to where you were before the jump, and from there you kind of open very carefully, you can open the design space again and explore sounds around that centre, and I think that was a very nice way of playing around because you capture things. I found that even more useful with the foam because it's so hard... it's harder to control and find things again'

²⁸Participant 2: 'I spend years controlling music, it feels a bit wrong to let the control be an algorithm and be randomly generated. I lose that sense of the sound is mine, that I am gesturing, that I am creating'

²⁹Participant 4: 'Often coming up with a sound is quite laborious. I've got a very basic knowledge of how to

for discovering unexpected sounds.³⁰ One person stated their preference for exploring larger spaces, describing some random settings as too limited.

The centering function was also discussed. Four participants commented on how this was a useful feature,³¹ although the jump in setting caused by the centering took some adjusting to³² and could interrupt the flow of exploration. Participant six commented on how they liked this delineation of flow.³³

Three participants talked about how they enjoyed the process of discovery afforded by the search system.³⁴ Participant three felt it compared favourably to interactive evolution.³⁵

The participants discussed the workflow they favoured when using RECZ. This typically involved zooming right in to a constrained setting and then moving outwards again, slowly opening up the space.³⁶ Constraining the search space in this way allowed subtle manipulation of the sound.³⁷

7.6.4.4 PLAYER EXPERIENCE

Five of the participants commented on how they enjoyed the freedom of motion available,³⁸ in particular with Phalanger. Using the foam, fewer degrees of freedom of motion were possible so this restricted the possibilities of expression in comparison to using the unconstrained hand.³⁹ Participant six considered Phalanger as freeform enough to express a personal style

adjust the basic parameters to get to a sound. Having that random element really would help me', participant 1: 'I think all settings are good for a start. Because when you start moving or touching you are going to get another different sound from the first one so, it's just a point to start, and if you don't like it you change it ... it's very easy, I like that. It's a great idea'

³⁰ Participant 4: 'The randomisation is a nice feature as well because it takes you somewhere you don't expect to be. Just by playing around you wouldn't yourself find the way there, but randomisation certainly is something that is in that respect very good'

³¹ Participant 5: 'I think one of the best things is that you can centre around something that you found interesting and I used that methodologically more and more often'

³² Participant 4: 'It took a little bit of getting used to, having to let go of where you were... you need to let go after you'd centered it? To get back to that.'

³³ Participant 6: 'In the beginning I didn't like it but I got used to it. It's like a full stop in a sentence. It's like, ok, stop you're moving on. In a way it helps you also to switch, if you switch your imagery or switch your mindset for a particular sound. Yeah but I don't know what it would be like if it was continuous'

³⁴ Participant 1: 'That is much more intuitive, you can find things that I will never think about', '...it became just a soundspace that I could explore. The discovery process was the fun part'

³⁵ Participant 3: 'I like the way its done a lot. I had thought about doing a similar thing with just randomisation and genetic algorithms and just doing audio feedback. It's a lot more fun moving your hand and playing the thing than like being like, ok next sound. With this you have control over it'

³⁶ Participant 5: 'I found a sound, centered it, then zoomed in so that I could reproduce the sound that I found ... and come out again and see what the space around that sound is. For me that was the means of exploration; find a sound that is interesting, centered it, zooming right in to that sound, had it reproduced and then went from there', participant 5: 'But when I found something I quickly centered it and zoomed right in, then I could let the foam expand and I had the sound still there, and then I could kind of open up the space again.', participant 6: 'I started somewhere and then slowly zoomed in. I'd zoom down to really close to see what the difference was and if I found a place that I thought was really good then I would stay there for a while and then I move around a bit until I wanted to move on', participant 6: 'I would go to a random preset and then I would say, ok that sounds quite holistic. I mean this is what I think I was thinking... let's see if this space has some areas which are interesting, and then I would hang out for a while, move around, find something I could centre on or not, and then if it was not that interesting I would move on to the next one'

³⁷ Participant 4: '[I noticed you were zooming right in quite close to...] Then you get those subtle tweaks, when the resonance is making a tone; I enjoy those kind of sounds'

³⁸ Participant 1: 'When you are using normal software, you are more constrained, you have not so many possibilities'

³⁹ Participant 6: 'With the foam I found that it was mostly pressing on areas or turning it in certain ways, so I found that this was more constrained, and hence I couldn't get the sense of sculpting'

of interaction.⁴⁰ Continuing the theme of personalisation, participant five discussed how they created their own parameter set of hand motions.⁴¹ Participant six talked of their sense of dialogue with Phalanger, again creating their own language of motion.⁴²

For participant five, the visuals available in Phalanger helped them to see the possibilities of different hand shapes they could use to interact with the system, and to see how the system picked them up.⁴³ There was more discussion of the use of visual reference while using Phalanger; it helped participant four to use the system⁴⁴ although they weren't completely sure why, while participant six was much less focused on the screen.⁴⁵

The feeling of embodiment when using the two interfaces was a key topic in the discussions. Participant six described the connectedness they felt when using Phalanger,⁴⁶ between the imagery in their head and the motion of their hand. With Phalanger, participants used their proprioceptive senses to explore the soundspace,⁴⁷ in contrast to exploring the material interface of the foam with touch,⁴⁸ the foam representing an abstraction of the sound. Half of the participants talked of the use of metaphor when using the system. Especially with Phalanger they talked of moulding⁴⁹ or sculpting⁵⁰ the sound. For this latter participant, this use of imagery and metaphor was part of a deep sense of engagement with Phalanger.⁵¹ They

⁴⁰ Participant 6: 'because it's so freeform it doesn't... it's part of the adventure rather than a constraint. It feels very free, very idiosyncratic. I think we have ways of moving, everyone does, and you're not constrained by anything, and that's what I liked about it'

⁴¹ Participant 5: 'I made my own parameter set by using the space I had with my hand, so that was like lifting it off the surface, doing different shapes. And they obviously didn't necessarily map onto the parameters directly but it was my own sense of the synth space. So I quite liked that.'

⁴² Participant 6: 'And I have to say that there is a sense of dialogue ... because of the dimensionality with the hand movement. Because then you'd start to create a certain hand movement which you could think of as the syntax of the interface or the dialogue that you're having'

⁴³ Participant 5: 'I looked all the time at the screen actually. ... I watched the screen I think mainly because I like the green hand, it looked funny. And I also got a sense of what the system actually picked up. Probably, looking at the representation rather than the hand itself, you become more aware of the shapes that you can produce. I wouldn't have thought about all the different shapes I could do if I'd just looked at my hand, but seeing the recording on the screen I think helped to become aware of that'

⁴⁴ Participant 4: 'To begin with it was easier to use than the squeeze thing but that's because there was some visual feedback it was giving me, even though it wasn't feedback that was actually saying what was going on. Something to focus on maybe, I'm not sure'

⁴⁵ Participant 6: 'most of the time I wasn't really aware of the visual so much. I was really in my head'

⁴⁶ Participant 6: 'So it was all much more connected, the imagery in my mind, the gesture in terms of making objects in space, caressing dimensional objects in space even though there weren't any, as well as the directing of the sound into a direction that felt it was something I want to explore'

⁴⁷ Participant 5: 'I think in the initial exploration, with the hand I kind of had more time to listen to the sound because my hand was doing... its my hand, I know pretty well what its doing right? I don't need to think about what my hand is doing'

⁴⁸ Participant 6: 'The adventure I was just telling you about, it's a more internal journey in that case with the camera, and this is more a focused journey so in other words with the foam I was looking at what the foam can do for me, I was putting that into my objective and the camera interface.. the objective became more let's see what it looks like where I'm going. Its not a hierarchical thing, it's not about better or worse, it's just two different things I feel'

⁴⁹ Participant : 'I came at it with the same idea where I might use my hands to sort of open and close my hands this area so the camera sees a rather large area, a block of space that I mould and in the same sense this foam is another block of space that happens to be smaller so I was kind of coming at it with the same sort of logic'

⁵⁰ Participant 6: 'I was really in my head, and in fact at one point I was thinking of it actually being more like a sculpturing process, I was sort of embracing the forms in the space... I had a dimension metaphor so I was thinking I could sculpt the sound. I had the sense that I could actually, when I made a movement, I was embracing, I was caressing something, it wasn't just air.'

⁵¹ Participant 6: 'The camera interface provokes kind of an entrancing quality, it's definitely a deep immersion which caters to all the different senses'

also imagined a metaphor while they used the search process.⁵² The foam controller, with fewer degrees of freedom, had much less possibility for the use of metaphor, or to provide a semantic link between the player's motions and the resulting sound.

The were comments on how the system could be used as an improvisational tool, and how using it for a performance would be a visual spectacle, although one person felt it might be too random for live performance.

7.6.4.5 METHODOLOGY

This is a brief category but worth noting. The following was said by a participant who had taken part in the previous EchoFoam evaluation: *'I liked the sound better than last time, so it was more fun to play around with.'* This is continued confirmation of a previously noted methodology problem with this style of study; the musical material can influence the players opinion of the controller being evaluated.

7.7 Discussion

Overall, the results were very positive about several aspects of the systems the participants tried during the study, especially concerning RECZ and the use of Phalanger in this scenario. The results highlighted some differences between the two interfaces, and demonstrated how an increased sense of embodiment underpinned the successful aspects of the systems.

The reaction to RECZ was very positive. It worked well as a sound discovery and exploration mechanism, and it worked well as a means of high-level control in a complex multiparametric system. One potential issue with the system was with the centering mechanism. Participants in the initial evaluation sessions found that centering the subsets caused a jump in the sound output, and the player would have to explore with the controller to get back to the original point they were at before the jump. In this study, the centering process was less of an issue; although the same effect still occurred, one participant saw it as a positive break in flow, and others bypassed it when they used a strategy of zooming far in and then slowly outwards. This style of use of the system was one of the more unexpected results of the study. The system was designed with the view of gradually zooming in towards a sound, instead participants tended to use the strategy of zooming in to near maximum and then scanning the space by zooming out and then in again at high zoom levels. At the higher zoom level, the effect of the jump when centering can become less noticeable.

Looking at the differences between the two controllers, participants were generally much more positive about their experience of using Phalanger than the EchoFoam controller. The key to this difference was the high degree of freedom of motion afforded by Phalanger, which was much higher than that afforded by the foam. This allowed the player a finer degree of control, but most importantly allowed the player to build a semantic relationship between

⁵²Participant 6: 'There was that sound and that sound and now you can go to that sound but you can change it slightly. You can explore a bit further. So it became this tree that you can go to, and I think it really works better with your imagination too in terms of visualising, because I know I had visual imagery before but it was just a more static imagination whereas this was more of a travelling kind of imagination.'

their hand motions and the sounds they were generating or exploring. While the more limited motion of the foam forced the player in certain directions, Phalanger's freeform interaction allowed their movements and imagination to roam freely while playing the instrument. This allowed metaphors to be built by the players, and with some participants this created a deeper sense of engagement, connectedness and dialog with the system.

7.8 Conclusions

Having analysed the results from both the informal and formal evaluations, some conclusions can be made about this system.

- Continuous multiparametric control can be a successful mechanism for sound exploration.

Some participants had an engaging and enjoyable experience using the two controllers for timbre space exploration.

- Participants were divided in their opinions concerning the holistic style of control.

Some participants valued holistic control, while others found this difficult. This adds weight to Hunt and Kirk's conclusion that some people prefer to think in terms of separable parameters, and therefore find multiparametric control challenging (Hunt and Kirk, 2000).

- The search system worked successfully.

Reactions were positive, indicating that this mechanism of high level control of the subset of explorable search space could be a valuable tool for musicians. In terms of interactivity, one possible factor for the system's success in this scenario was the use of bimanual control (Wilson, 1998, chap. 8). The players could use one hand for detailed control with Phalanger or the EchoFoam, while using the other for higher level exploration, controlling the search engine with a slider and keys. This builds on our natural mode of interaction when using bimanual systems (Treadaway, 2009).

- Intangibility can be a strength as well as a weakness of a controller.

Previous studies with Phalanger showed that the lack of physical resistance in the interface could be a problem for some users. However, in this scenario the lack of tangibility was a definite advantage. It allowed freedom of motion, and in turn allowed better sound-motion correspondence and an increased engagement with the system. Speculatively, could the use of metaphor compensate for the lack of physical resistance? Does metaphor create imagined resistance, in some way making the system more 'tangible'?

- In systems with a large degree of freedom of motion, a choreographed approach to control can benefit the players experience.

Some participants created a language of control to use with the system, and felt that they benefited from it. When using a complex multidimensional system, creating a choreographed language of motions is one way of delineating the large space the user is working within. It can also help the player with repeatability, and with memorising larger structures.

CHAPTER 8

DISCUSSION

‘I think that electronic technology offers us the possibility of divorcing ourselves from the necessity of virtuosity, without divorcing ourselves from the possibility of intense and meaningful interaction with our instruments.’

Don Buchla (Diliberto, 1983)

8.1 Looking Back

Before starting to make conclusions from the studies carried out for this thesis, the story that brings us to this point is briefly summarised. At the beginning a research area was set out which focused on musicians interacting with standard computers to create music. It was established that some difficulties for computer musicians using tools designed for production work are related to the low quality of embodied interaction that the systems afford, through limited bandwidth input devices.

A design space was set out, listing desirable qualities of new interfaces for GUI based digital music tools. Based on this, two new interfaces, Phalanger and EchoFoam, were designed. These interfaces were created principally with the aim of affording a high quality of embodied interaction, exploiting the perceptual-motor skills of the user. Both interfaces were evaluated formatively in laboratory based studies before being tested in field studies. As a study in further engaging with multiparametric interaction, an interactive system for exploring complex timbre spaces was designed, named RECZ. This system could use either controller as an input device, and was evaluated with both in a laboratory based study.

Having described and discussed the results of these five individual studies, I would like to reflect on the combined results in order to elicit conclusions about both controllers individually. I will then analyse the combined results to reach some more general conclusions about the research questions set out at the beginning of the thesis. I will start by reviewing the methodology used in this thesis, both to frame the validity of the results and to provide critical feedback for those considering using these methods themselves for the evaluation of digital music controllers.

8.2 Review of Evaluation Methodology

In chapter 4, a methodology was set out for evaluating the systems built in this thesis. The unique nature of the evaluation of digital musical controls was discussed; this is a novel and difficult problem for HCI, and there is no standard, well tested technique for evaluating player experience. The methodology I have used in this thesis is a conglomeration of evaluation and analysis techniques drawn from HCI and sociology, adapted for a specialised and complex problem. Given the novelty of these techniques applied to computer music, I would like to consider their effectiveness, and look at how they could be improved both on conceptual and practical levels.

8.2.1 FORMATIVE EVALUATION METHODS

The case study of evaluating the Wiimote as a musical controller (in section 4.2) sets out the issues surrounding the use of laboratory-based studies for gaining insight into the use of musical controllers. Three more formative evaluations like this have been carried out for this thesis. All were based on the methodology adapted from lessons learnt running the Wiimote study; all the studies involved participants carrying out pre-designed musical tasks with new controllers,

using preset musical materials. Data from the three studies came from post-task interviews with the participants, and all the data was analysed using a grounded theory approach. All three studies focused on player experience, and were purely qualitative. Reflecting on these studies, there are three further principal areas where weaknesses were apparent. These issues are additional to those already mentioned in the reflections in section 4.2, on practice time, variations in skill levels and unnatural musical settings.

- **Aesthetic Preferences.** Given the short length of the studies, all musical materials were set up beforehand, and were identical across all participants to preserve consistency. For example, in the formative evaluation of the EchoFoam system, all participants tried the same phase modulation patch. An issue that became apparent, and which participants commented on in two of the three studies, was that aesthetic preference concerning the musical materials may affect the participants' view of the controller. There is no easy answer to this; there is no such thing as neutral musical material. It would be possible to allow participants to create their own patches in a study, but this would bring additional complexity, and distract from the main focus of the study.
- **Novelty Factor.** When a new controller or a new technology is used for the first time, there may be a novelty or 'wow' factor which lends the controller a more positive light than is actually the case, the positivity being directed towards the technology rather than practical aspects of use. For example, it might be fascinating and exciting to use a hand tracking system for the first time, and this excitement may colour the participant's opinion of faults in the system. This 'wow' factor is something that would naturally wear off over time, but not necessarily within the period of a short study.
- **Confidence.** Not all participants were clearly comfortable with exploring a controller for the first time while sitting with a researcher. This is especially the case in studies with non-musician participants who have little or no performance experience. This confidence factor means that some participants will be shy about exploring an controller to their full ability, and will not be able to give high quality insights compared to someone who does not feel they have to hold back in front of the researcher.

Focusing on the positive outcomes, all three studies yielded valuable insights into the controllers they evaluated, and highlighted important issues surrounding their use. These insights were largely beyond the intuition of the researcher, making the results very worthwhile. For the two studies that preceded field studies, these laboratory-based evaluations revealed key usability issues that needed to be resolved before sending the controller out for more in depth evaluation. Had these issues not been recognised, they would have affected the quality of results in the studies that followed.

Overall, there are pronounced issues when using this style of controlled study, and most issues are rooted in the way these studies take place in contrived scenarios. Having said this, there is no doubt that this kind of study is extremely useful, especially in the early stages of design. All results of these style of evaluations must be considered in the context of these methodological problems. Practically, a balance must be considered between the need to obtain fast early results which are critical for steering the design of a system, against obtaining truer reflections on an instrument from observing its use in a more natural environment over a longer time period.

8.2.2 FIELD EVALUATION METHODS

In section 4.3.3 a case was set out for the use of field evaluation to evaluate musical controllers. The principal reason for this was to try and obtain a truer picture of musical interaction that can only take place outside of a controlled setting. In doing this, usage of an instrument is not limited to the designers' intentions, and creativity is encouraged in a natural environment. This also means usage is socially situated, and allows the player to gain skill and experience of using an interface over time.

Both EchoFoam and Phalanger were evaluated using field studies. Reflecting on these studies, there were a number of issues concerning this style of study:

- **Participant Motivation.** In both field studies, participants were asked to collect data about their experience of using the controllers such as diaries, photos and audio clips. While some participants put effort into doing this, others were much less diligent in this respect. Motivating participants to keep records is a difficult issue, as the researcher must be careful to keep a balance between gathering more information and interfering in a study in a way which will colour the results. Further to this, gathering too small an amount of data can also colour the results. In the Phalanger evaluation, participants were unpaid volunteers, and in the EchoFoam evaluation, participants received the small reward of a free controller and software, so perhaps a greater payment for participation may have motivated participants. One strategy would have been to require submission of regular diary reports, though this goes against the aim of allowing the participants to use the controller in a freeform way, and unnaturally restricts the format of data being collected. It would also be difficult to enforce for unwilling participants. Another strategy could be regular phone interviews but again this could be intrusive to participants.
- **Usability.** Software and hardware made for a field study needs to be robust enough to work effectively outside of laboratory conditions; usability issues will colour the results. Achieving this level of quality can be difficult as the system being tested is typically in a development stage, made by a small development team, and using experimental techniques for interaction. While there were few of these issues in the EchoFoam study, there were a number of difficulties with the Phalanger system, mostly concerning lighting and skin detection. Sometimes usability problems are difficult to pick up until a system goes out for real-world testing. In running this style of study, issues will invariably reveal themselves, so possibly the best strategy is to be ready to resolve any issues as soon as possible and send software or hardware updates to the participants. In this way, usability issues will have a minimal effect on evaluating user experience.
- **Open Endedness.**

Both field studies were designed to be as open ended as possible. The software was built to allow completely freeform use of the interfaces. Phalanger was completely configurable; users could train it to respond to any set of hand poses and map it to any parameters in their music generation software. EchoFoam allowed a number of mapping configuration options in the software, and could be connected to any continuous parameters in Live. This meant that participants could explore the interfaces and software in any way which they preferred, and fit the interfaces into their own working practice. Limiting the interfaces to perform a contrived task would contravene the aims of a field study of a creative tool. The difficulty with this approach is that the participants can follow very different paths and have very different experiences with a controller. This can yield good results because it highlights different uses of a system, though it also means that there may not be much overlap between individual participants' experiences. As a

result it may be difficult to find common threads that can be cross-referenced between individual results. This may leave a small number of strong threads from which to make conclusions. It follows there may also be a larger amount of weaker threads that seem compelling but do not corroborate well between users.

- Time. This style of field study is very time consuming to run. Software and hardware must be developed that is robust enough to survive real-world usage and abuse by participants. Following this, participants must be found, trained in the use of the system and interviewed after and possibly during the study. Data must then be compiled, transcribed and analysed. This all adds up to a large commitment in time and energy, but the reward is a set of detailed results.

8.2.3 COMBINED METHODS

In evaluation of the two controllers, both formative laboratory-based and field based evaluations were carried out. It was intended that these two styles of study would complement each other. The field study would compensate for the methodological shortcomings of the formative evaluation, moving from a contrived scenario to the very natural scenario of the artist in their own environment. Further to this, the formative evaluation would be a useful way to obtain early results, and to help orient the field study to obtain the best results. In the aftermath of the two studies, the success of this combination can be examined.

It's evident that the formative studies were a valuable precursor to the field studies. For both interfaces, the formative evaluation helped to iron out bugs and identify important functional issues concerning the controllers' usage. For example the formative evaluation of Phalanger identified ergonomic and performance issues that were corrected before the field study. The field studies revealed information about usage that couldn't be obtained in the controlled setting of the formative evaluations. For example, the EchoFoam field study saw the controller being used for a variety of different musical styles, by artists in their own studios, and revealed how the controller fit into peoples' typical creative process.

Overall, it seems that the field studies needed the input from the laboratory-based studies to guide them in the right direction, whilst the results from a laboratory-based evaluation aren't enough on their own to get a wider picture of a controller. These two methods have complemented each other well, each one needing the other.

8.2.4 EVALUATION CONCEPTS

The data from the interviews conducted for this thesis was analysed using a grounded theory approach. One of the outputs of this is a set of categories which emerged from the analysis of participants' responses, illustrating the concepts around which the data was clustered. This collection of concepts may prove to be useful to others designing controllers or running similar studies. For designers, it highlights issues which may be important to the users of multiparametric controllers. For an evaluation study, it provides a list of framework around which an interviewer may wish to structure an interview session. The word cloud also provides a compact summary of the combined analysis of the studies.

These concepts are illustrated in a word cloud in figure 8.1 and are also listed as text in appendix C. In the word cloud, more frequently occurring concepts appear in a larger font.

To create this list and word cloud, the concepts were collected from the grounded theory analysis of the interviews in all five evaluations of Phalanger and the EchoFoam system. They were then sorted alphabetically and normalised, using the following rules:

1. Concepts that were very specific to one controller were generalised if possible. For example, ‘comparison-to-old-foam’ was changed to ‘comparison’.
2. Concepts that were too specific to generalise were removed.
3. Concepts with identical meanings but different names were merged.

Duplicates were preserved in the list, to indicate the frequency of occurrence.

8.3 Combined Results

The two interfaces created for this thesis have been evaluated with two studies, and were also contrasted in the timbre space exploration study. This has yielded three sets of evaluation results for each system. In this section, I will triangulate from these studies for each controller, to build a more complete picture of each system; their strengths, weaknesses and overall success. This will be done within the framework of the design space set out in section 3.5, the evaluation results revealing the level of success of the controllers in each category.

8.3.1 PHALANGER

8.3.1.1 CONVERGENT INTERACTION MODEL

Phalanger’s design fits in very well to the convergent model of interaction. It takes input from the movement of an unconstrained hand, which gives the system an extremely high EDFM. This motion is tracked with a camera, and typically reduced to ten parameter streams which describe the high level geometry of the hand. A further discrete parameter stream is available for describing the hand shape as recognised by a trained SVM, although this parameter was not used in the timbre space navigation study. These inferred parameters can then be mapped to any musical parameters or further mapping algorithms.

This large freedom of motion was simultaneously the most positive aspect of the system, and a cause of a number of difficulties in interaction. Focusing first on positive aspects, users in the field study found the system to be very expressive, as did participants in the RECZ study, who appreciated the freedom of motion possible compared to the foam controller. The method by which the hand was tracked and reduced to a small amount of parameters meant that gestures could be repeated in varied ways to achieve the same effect. In turn this meant that users could employ ancillary gestures, bodily gestures that do not generate sound but form part of the expressiveness of playing an instrument. Ancillary gestures can be seen as an ‘integral part of musical performance’ (Wanderley, Vinesa, Middletona, McKay, and Hatch, 2005), and accommodating this type of motion increases the potential for musical expressiveness. The freedom of motion afforded by the system also allowed users to discover a personal style of interaction, something that was commented on in both the field study and the RECZ study. This personal style was revealed in how users employed different gestures

to play the controller, and also in the way they could configure their own set of hand positions to communicate with the system.

The large freedom of motion also created some difficulties. One participant in the field study found that creating a language of motions to communicate with the system was problematic because it was difficult to imagine how to configure the system. The large freedom of motion also meant that it was possible to create gestural mappings that have no correspondence to the sonic output, and this lack of symbolic connection could detract from the experience of using the system. A significant factor that influenced participants' view of the system was imprecision, and this was closely related to the freedom of motion; with no physical feedback users relied on proprioceptive and visual senses to operate the system. Some users in all three studies had no problem with this, and in fact found the system to be accurate; other users found this more challenging, and felt the system lacked accuracy and repeatability.

Overall, the convergent model embodied within Phalanger was a successful aspect of the system, but it also added some significant complexities to interacting with the system, which some users felt uncomfortable with.

8.3.1.2 OPEN DESIGN

There are two important factors in openness of design. Firstly comes the potential for use in a wide range of scenarios, acknowledging that the system will be used outside of the intentions of the designer. The other factor is fault tolerance, making sure the system is robust in a wide range of situations and environments. Phalanger seems to have the potential to be used in a large range of musical scenarios. The laboratory-based evaluation demonstrated the system in a range of situations with varying types of control. Following from this, participants in the field study took varying approaches to the system. The RECZ study demonstrated another application for timbre space exploration, and further showed that it could in theory be used by any system which takes multiparametric input. The system however is lacking in terms of fault tolerance, and the principal cause of this is oversensitivity to lighting changes in the environments. While this system works robustly in a constant lighting environment, changing lighting causes errors in the skin tracking which propagate through the system. This gives the system a lack of portability, and means it would be difficult to use in performance with stage lighting.

8.3.1.3 LEARNING CURVE

The laboratory-based study showed a mix of initial reactions to trying out the system; some participants quickly adjusted and others found it difficult to use, partly due to its unconventional nature. This study however only reveals data about an initial short period of use. The field study revealed that in the long term, the learning curve was more manageable. Users became steadily more competent over time, at both playing the controller and at training it with new hand positions.

8.3.1.4 CONTINUOUS CONTROL

All control in Phalanger is continuous apart from the discrete output from the hand position classification system. This gives the system the potential for fluid operation of an interface or sound engine, though the nature of this operation is largely dependent on the mappings the user has created.

8.3.1.5 INTUITIVE CONTROL

Users in both the field study and timbre space study found Phalanger to be intuitive to use. For example, one comment was that it was ‘instinctive and natural’. The field study revealed that intuitive feeling in use was enhanced when sonic output corresponded symbolically with the players’ hand motion

8.3.1.6 FINE-GRAINED CONTROL

Some users of Phalanger did not feel that the system was suitable for fine-grained control, feeling that it lacked accuracy and repeatability. Other users however felt the opposite, and the reason for this cannot yet be explained. In the timbre space study, one participant talked of making both broad and detailed changes in the sound, also finding they were able to reproduce gestures. In the field study, another user talked about a close bodily connection with the music, and described the expressive potential as similar to an acoustic instrument. From this evidence, it seems that Phalanger has the potential for fine-grained, expressive control, but this may be the feeling of a minority of users.

8.3.1.7 UNCONTROL

Uncontrol concerns the non-linear, unpredictable behaviours of a musical system, and their balance with linear, predictable behaviours. The results from the field study showed that a degree of uncontrol was inherent in the system, but also showed that this could be a factor people enjoyed when playing the system; it allowed freedom for error in a similar way to an acoustic instrument. The timbre space study showed that the level of uncontrol could be reduced when used in conjunction with RECZ, which allowed interactive control of constraints to the parameter space. Overall it seems that, dependent on the chosen mappings, there is potential to manage the degree of uncontrol. Unpredictability will always be present in the system due to the intangible nature of its use, but this can be optimised through careful configuration. Further to this, for some users, uncontrol enhanced their experience of using the system.

8.3.1.8 SUMMARY

In summary, the combined results show that Phalanger meets the aims of the design framework well; it affords the necessary types of control, has a manageable learning curve, has a convergent model of interaction, and can be used in many varied musical scenarios. The area where it falls down is fault tolerance, where it becomes unstable in variable lighting environments. This is a key issue that needs to be solved.

8.3.2 ECHOFOAM

8.3.2.1 CONVERGENT INTERACTION MODEL

The EchoFoam controller has a high EDFM; it takes a very large set of motions as input, which includes all the movements a user can employ to manipulate malleable foam with one or both hands. The controller outputs a twenty-five dimensional stream of data describing the shape of the foam, and this data is reduced in dimensionality, typically to between one and ten control streams that are mapped onto musical parameters. Being a tangible system, the EchoFoam controller is more constrained in use than Phalanger. This was evident in the timbre space study, where participants almost universally preferred Phalanger's larger freedom of motion. Further to this, EchoFoam users have less freedom to create metaphors, connecting actions with sounds. In all three studies, participants were divided into two camps that either enjoyed the high level, global control afforded by this model, or preferred the separable parameters in conventional controllers.

The tangible medium through which motion was captured caused some problems. The foam material suffered from latencies due to its slow expansion rate. Some also found it hard to operate, feeling that repeatability and accuracy were difficult, and that there wasn't enough physical feedback. Overall, the convergent model in this system was the root of its strengths, but also for some users it caused some significant problems in interaction.

8.3.2.2 OPEN DESIGN

In terms of potential for varied applications, the studies demonstrate EchoFoam being used in a reasonably wide range of scenarios. Principally, it was used as an instrument for the exploration of timbre spaces. In addition, the field study showed the controller being used for detailed editing, and collaboratively in a group setting. In terms of musical material, participants universally found it to be more useful for continuous sounds rather than percussive sounds, so this is one limitation in openness. Although users have explored and discussed its potential as a performance tool, it may prove more suited to studio work.

The system has shown itself to be highly fault tolerant. In the field study there were only a few minor software issues. Further to this, no participants experienced any hardware problems, so it seems as if the system is very robust.

8.3.2.3 LEARNING CURVE

Both the laboratory-based and field studies showed a mix of learning curves. The results from the laboratory study are limited because of the time spent using the system, though it showed the initial reactions of users, some of whom found the controller difficult to pick up and some of whom adapted easily. The field study revealed that while most users found the controller easy initially, some found progression difficult. This was related to controllability issues, and skill needed to control the mappings system effectively.

8.3.2.4 CONTINUOUS CONTROL

All control in the EchoFoam system is continuous, both in the physical controller and in the software. This means that interaction can never take place in 'analytic mode' (see section 2.3), all control is performative.

8.3.2.5 INTUITIVE CONTROL

Users in all the studies found the EchoFoam controller to be intuitive to use. People talked of having a direct, tactile connection to the sound, and also believed that the act of manipulating foam felt natural.

8.3.2.6 FINE-GRAINED CONTROL

In the initial study, most participants felt that high accuracy would be difficult with the controller. There were mixed reactions about repeatability, some feeling it was reasonable and others finding the foam material complex to manipulate and perform the same motions twice. Before the field study and the RECZ study, the sensitivity of the controller was greatly enhanced, and RECZ added interactive control of the parameter space. With these factors changed, there was still a group of users who found the controller difficult to use accurately. However, some users also found the controller to be accurate enough for detailed work. The key to this was the parameter space system; with carefully configured parameters the controller could achieve a good level of accuracy. This level of accuracy however did not compare to conventional MIDI devices because of the nonlinear nature of the controller and mappings; with codependent multiparametric control, there is always a balancing act happening between altering parameters.

8.3.2.7 UNCONTROL

Uncontrol was one of the most important aspects of the EchoFoam system, and this was highlighted in all studies, where participants used it for discovering new and unexpected sounds. Uncontrol is made explicit in the EchoFoam system; the control streams are codependent and nonlinear which means that for each new mapping, the user is forced to explore how the controller works in that particular configuration. This means that as mappings change, there will always be new surprises for the user. Participants in the studies enjoyed using the controller discovering unexpected new sounds, and for creative inspiration. Uncontrol was also the biggest hindrance; some users did enjoy this kind of behaviour in a controller, and some found it too difficult to use, perceiving it as being too random.

8.3.2.8 SUMMARY

The evaluation results show that the EchoFoam system has fulfilled the general aims of the design framework, but with some prominent problems. Firstly, the potential applications could have been wider, a key problem to address is how to map this system to more percussive sounds. Another issue is the way in which the system captures motion: the latencies in the foam have caused problems, and some people found the malleable material difficult to

use for musical control. This seems to be related to user preference; while some users have very much enjoyed and adapted to using the material, others found it difficult or plainly did not like it. Lastly, the degree of uncontrol proved to be very high with some mappings, and some users felt the system was too unpredictable. Participants in the studies have had strong preferences concerning this system; in general while some had a rewarding experience, others felt uncomfortable with it.

8.4 Emerging Themes

Being that both Phalanger and EchoFoam have been created within the same design framework, they both share some key properties: they are complex, multiparametric, nonlinear interfaces with codependent control streams, that emphasise an embodied approach to interaction with digital music systems. There were several strong themes which emerged from the combined results of the evaluations of both systems, which yield some more general insights into the use of this class of system.

8.4.1 METAPHOR, IMAGERY AND SOUND-MOTION CORRESPONDENCE

The Phalanger field study emphasised the importance of metaphor and imagery in creating successful, intuitive mappings between motion and sound. Looking at the results from the other evaluations, this issue is highlighted further, and is emphasised by the differences between Phalanger and EchoFoam. In the EchoFoam results, participants in the formative evaluation talked of satisfaction when a sound-motion correspondence existed. In the field study, there were no reports of any participants using metaphor or imagery when using the foam. This is in stark contrast to the Phalanger system where these were important concepts concerning the system's use. The RECZ study highlighted the contrast; participants talked of the freedom that hand motion gave them to create metaphorical links between their actions and sound, with people talking of sculpting and moulding the sound, and describing some vivid multisensory imagery during use. They also reported that EchoFoam felt restrictive in comparison, and did not allow the freedom for metaphor in the mappings.

Three interrelated concepts concerning this area can be seen in the results: metaphor, imagery and sound-motion correspondence. Participants felt enjoyment and satisfaction when sound-motion correspondence existed in mappings. With Phalanger they employed metaphor and imagery to create the symbolic link between motion and sound; the results also show that when this link doesn't exist, it can detract from the experience of using the interface. One can look to Jensenius' work to find some insights into these results (Jensenius, 2007).

Jensenius focuses on the relationship between action and sound. He stresses the strong perceptual relationship between them, and talks of two types of connection, action-sound coupling and action-sound relationship. Couplings are based on real-world experience, while relationships are artificially created connections for artificial sounds such as those generated electronically. Jensenius proposes that to construct a good action-sound relationship, it should be modelled on real-world action-sound coupling, and stresses that this is an important factor when designing mappings in a musical controller.

Looking at the two controllers through this lens, it needs to be emphasised that both systems have very fluid mappings. With EchoFoam, the mappings are often changing and explicitly need to be explored, and in both systems the mappings are very configurable by the user. When using RECZ, the mappings are in continual flux as the user explores the sound space. Further to this, as these are multiparametric controllers, when one mapping changes, the entire control landscape changes. This change may be dramatic even with small mapping changes due to non-linearities in the mapping algorithms. As these mappings change, it is the responsibility of the user to try and create any artificial action-sound couplings. As Jensenius says, the best couplings are based on real-world examples, and this provides a good possible explanation of why users preferred using Phalanger for timbre-space exploration. Given the large freedom of motion available with the system, it allowed the users to create their own perceptually matched way of interacting with the sound space, and gave them freedom to use metaphor and visual imagery to create the symbolic link between motion and sound in an abstract space. It is also a possible reason as to why playing the interface was more difficult when this symbolic link could not be made. EchoFoam on the other hand, provides a pre-set metaphor of malleable manipulation, and does not give the player such freedom to make metaphorical associations.

This shows that the capacity for metaphor and imagery can be an important part of controller design, especially where mappings are in flux, as is usually the case in digital music where the user controls the mappings. It highlights the potential that intangible control elements have for allowing the user to realise metaphor and imagery, and therefore create perceptually matched action-sound relationships. How different would the experience of the EchoFoam controller have been if it contained a small accelerometer to make an additional stream of intangible motion based control?

8.4.2 CHOREOGRAPHY AND LANGUAGE CREATION

Given the large EDFM of both controllers, the creation of a language of motions with which to use the controllers was an important issue for some participants in the evaluations. Conventional scientific instrumentation style controllers show their controls up front, and modes of interaction can be derived from visual inspection. These controllers afford a large number of ways in which they can be manipulated that were not obvious to the user without exploring such options themselves. The results demonstrate two types of language creation, namely explicit creation through configuration and language creation as a way of exploring and subdividing a parameter space.

In the Phalanger field study, language creation was explicit as the user could configure the system to respond to a customised set of hand positions. This meant that training the system became a skill in itself, involving creativity to define how the system would interact with the user, and requiring choreography and performance skills to create the sequence of gestures needed to effectively train the machine learning system. From my own experience of training Phalanger, I found myself creating detailed sequences of gestures to train the system for a particular scenario. I would then repeat these sequences with subtle differences to retrain the system for varied outcomes. To this end, I made an informal notation system to keep track

of the sequences while I trained the system. As I spent more time using Phalanger, acting out these gesture sequences became more intuitive; it felt as though training the program was as much of a skill as using the configured system. Each training sequence became a small choreographed performance, which I needed to perform well in order to achieve the desired outcome. This style of interaction with machine learning systems is a relatively new issue for HCI, and is explored in depth by Fiebrink, who created ‘The Wekinator’, an interactive machine learning tool, and studied musicians who applied it to digital music systems (Fiebrink, 2011). This topic is also addressed by Ashbrook and Starner (2010); they recognised that designers frequently have little expertise in machine learning, and created ‘MAGIC’, a tool to design and test gestures.

While training machine learning tools is one aspect of language creation for these controllers, the other important aspect concerns how players subdivide a complex and abstract multiparametric space into their own language of smaller motions. The evaluation results showed users of both systems discussing how they created their own language, more so with Phalanger where there was more freedom of motion, one participant describing these motions as the syntax of communication with the system. Participants in the EchoFoam field study talked of the vocabulary of motions they used with the system, describing them in terms of ‘squashes’, ‘twists’, ‘crushes’ and more. They also talked of how they mapped the cube into sections such as corners, so the vocabulary may have been a combination of motion descriptions and cube locations. With Phalanger, this language was described more in terms of metaphor, such as using a DJ deck.

To summarise, for the designers and users of multiparametric controllers such as the ones in this thesis, language and choreography are important elements to consider for improving user experience. Where the controller presents an large, abstract, explorable parameter space to the user, it’s important to employ strategies to make sense of this area. It’s also important to understand that a better language will create better symbolic links between motion and sound.

8.4.3 COGNITIVE STYLES AND MULTIPARAMETRIC CONTROL

Looking at the evaluation results as a whole, there has been a marked range of reactions to the controllers from the participants in each study. These reactions have ranged from enthusiasm to the systems in some individuals to dislike in others. In no study was there universal approval for a controller. In terms of results this is quite acceptable as universal approval was not the aim of the evaluations, but rather to understand the use and strengths and weaknesses of the systems. However, the interesting theme here is the wide spread of reactions. The results show this range of reactions in two areas: first the participants’ overall opinion of the controllers, and second, their initial ability to use the controllers effectively. As noted in the formative study of Phalanger, some people felt instantly comfortable with the system while others found it more difficult to use at first. It seems that participants’ reactions were divided along two sets of poles concerning control style preference and general musical approach. In terms of control style preference, participants were divided between linear and multiparametric control. Some disliked the inseparability and codependence of parameters

of the multiparametric approach. Others enjoyed the more holistic approach, and felt that conventional MIDI controllers were more limited in comparison. The participants' approach to music creation was where other divisions appeared, some preferring a more exploratory, emergent approach, and others preferring a top-down approach. This was particularly evident in the EchoFoam field study, where the controller almost forced an emergent approach.

There are a number of factors which may explain the breadth of reactions. A few participants commented on how they felt established in the way they used equipment and were less willing to adapt to a new approach. Further to this, the overall novelty of these multiparametric controllers may have been a factor, as participants were very unlikely to have used a similar system before. Beyond this, many participants felt uncomfortable with the lower precision levels in these controllers, compared to the highly accurate controllers they were accustomed to. A more universal explanation for these differences may be rooted in cognitive styles.

A cognitive style describes an individual's way of processing information (Steinberg, 2001), their preferred and habitual approach to organising and representing information (Riding and Rayner, 1998). Cognitive styles are typically described along bipolar axes, for example Pask proposes the holist-serialist dimension. Holists take a global approach to a task, focusing on the broader conceptual overview while serialists focus on small units, building up to the bigger picture which emerges later in the process (Ford, Wilson, Foster, Ellis, and Spink, 2002). There are several propositions of dimensions along which to label cognitive styles, Riding and Rayner group them broadly along two axes: wholist-analytic and verbal-imager. The wholist-analytic axis, analogous to Pask's holist-serialist dimension, describes whether an individual has a tendency to organise information structurally into wholes or parts. The verbal-imager axis shows how an individual tends to represent information when thinking, either as imagery or verbally (Riding, 2001). An individual's overall cognitive style is their position along these two dimensions. Their cognitive style influences behaviours such as learning, motor skills and social behaviours.

Eaglestone et al. (2008) found that the way composers approached composition tasks mapped well onto the global-analytic axis. Within their sample, the predominantly dominant global style composers tended to establish a structure and refine it, while the analytic style individuals tended to take an emergent approach to the process, building a piece from small components. Gelineck and Serafin (2009b) also found two similar global modes of composition, which they call *the worker* and *the explorer*. These dimensions map well onto the approaches of the participants in the EchoFoam field study, where the EchoFoam system seemed to work more successfully for those who tended towards an overall emergent, analytic approach to music creation. While this maps well onto compositional approach, it's not so obvious how cognitive styles might correlate with control style preferences. Speculatively, since parameter codependency and non-linearity lead to a level of unpredictability in an interface, a more explorational approach to learning may be required, making multiparametric interfaces more suited to analytic individuals. Conversely, the higher accuracy and separability of MIDI controllers may be more suited to a global/refinement approach.

A recent user study on individual interactivity preferences explores this area further. Kim, Lim, and Suk (2011) investigated different attributes of activity including continuity, concur-

rency and qualities of movement; they found that individuals have diverse preferences for some of these attributes, and that these were related to individual differences. Interestingly, they discovered that their participants expressed preferences for *concurrent vs sequential* and *continuous vs discrete* interaction. Furthermore, they found that preferences were diverse concerning concurrency; this echoes the results here, which show a range of preferences concerning holistic control vs separable parameters. They also found that 70% of participants preferred continuous to discrete control, which adds weight to the general approach of these controllers. Kim et. al's study was conducted using mouse control of a GUI, it would be fascinating to explore this area further in the context of digital music controllers.

8.4.4 UNCONTROL

Uncontrol has been a prominent theme throughout the evaluation results. Uncontrol (initially discussed in section 3.5) describes the unpredictable behaviour in a system, and encompasses the concepts of nonlinearity, randomness, unexpectedness, repeatability and precision in the control of a musical instrument. Uncontrol was an element in the design space set out in chapter 3.

Looking through the evaluation results, issues surrounding uncontrol occur frequently, and were a key factor in the usage of the two controllers. Users of Phalanger in the formative evaluation found that fine control was difficult, and felt the interface was suited to less accurate types of control. Others commented on elements of unpredictability in the system. This was emphasised further in the field study, which showed how uncontrol exists in a delicate balance in the instrument. Freedom of movement was a valued part of interacting with Phalanger, but this also meant freedom to make errors of control; this created a trade off between the quality of embodied control and uncontrollability.

The EchoFoam system was shown to exhibit uncontrol in a number of areas of its functionality. In fact, uncontrol was one of its most defining qualities for the participants in the evaluations. The system exhibited highly non-linear behaviour, and this was coupled with complex behaviour of the foam to create a large degree of unpredictability in the system. This level of unpredictability led to one of the most positive uses of the system as a tool for the discovery of new sounds.

RECZ showed itself to be very effective at damping the level of uncontrol in both controllers. Results showed that with carefully selected mappings, a higher level of detail and precision could be achieved. For example, in the EchoFoam field study, one participant felt the system was precise enough for detailed work. They also took a considered approach to mapping which contributed to this level of precision.

The results reveal some interesting insights into the nature of uncontrol. Before exploring these in more detail, I will briefly review what current research tells us about uncontrol in digital music systems.

Jordà highlights the interconnectedness of non-linearity with uncontrol. Non-linearity is seen as a source of uncontrol, but it can also provide more powerful higher-order control (Jordà, 2005). He believes that non-linearity should not mean complete uncontrollability or unpredictability; the performer needs trust in an instrument, and should be able to pull it back

from non-linear behaviour when needed. Gelineck and Serafin's (2009b) survey of computer music composers highlights unpredictability as an important factor in the design of musical systems; the participants enjoyed the surprises and accidents resulting from less predictable elements of the tools they used, and the majority felt that *'too much control can kill the creative process'*. Magnusson and Mendieta's (2007) survey of digital musicians revealed similar attitudes, one of the respondents is quoted as saying *'full control is not interesting for experimentation, but lack of control is not useful for composition.'* Bertelsen et. al's case study of two digital music composers shows how lack of control is an important element in the composition process. One composer talks of how his Max/MSP patches could reach a level of complexity which is difficult to follow, leading to unpredictable behaviour and creative inspiration. The other talks of the 'glitch' aesthetic of using errors and accidents as source material (Bertelsen et al., 2007; Bertelsen, Breinbjerg, and Pold, 2009). These studies present a view of uncontrol as something that is welcomed by composers as a source of creative inspiration, and as an important element in skilled control of an instrument. The evaluation results add to these studies to build a more detailed picture, presented here as a mini-taxonomy of uncontrol.

8.4.4.1 A MINI-TAXONOMY OF UNCONTROL

Figure 8.2 summarises the concepts surrounding uncontrol. Two groups of sources of uncontrol are evident, physical and virtual. Physical sources comprise uncontrol elements present materially in the controller, or due to the manner in which the user must interact with the controller. With the EchoFoam controller, the foam itself contributed to uncontrol because the material was always trying to decompress back to its default state, influencing the control streams as it did this. The foam itself is a complex nonlinear system, which is manipulated by the user to control sound. With Phalanger, the intangible nature of the interface was the principal source of uncontrol; freedom of movement led to freedom for player error. Virtual sources of uncontrol arise from mappings and algorithms processing the control data. For the controllers in this thesis, the ESN mapping method introduced elements of uncontrol because of the non-linear nature of the mappings, coupled with a slight inherent instability in the ESN outputs. For both interfaces, further elements of virtual uncontrol could be added by the player when the control streams were mapped to sound parameters. The EchoFoam study showed how important the choice of mappings was in achieving a reasonably stable, playable system; the system could easily lose too much stability given the wrong choices.

The causes of the existence of uncontrol fall into two groups: explicit and implicit. Explicit uncontrol is a deliberate uncontrol element introduced by either the designer or the user of the system. For example, the EchoFoam Live Link software allowed the random selection of mapping destinations and ranges by the player; the player could deliberately inject unpredictability into the creative process when looking for creative inspiration. Implicit uncontrol exists as a side effect of the principal design aims of a controller, for example output instability from an ESN or variability in hand motion.

The evaluation results reveal a little about perception of uncontrol. Interestingly, several participants in the EchoFoam studies perceived unpredictability in the controller output as randomness, when it was in fact due to the nonlinear and complex nature of the system. This

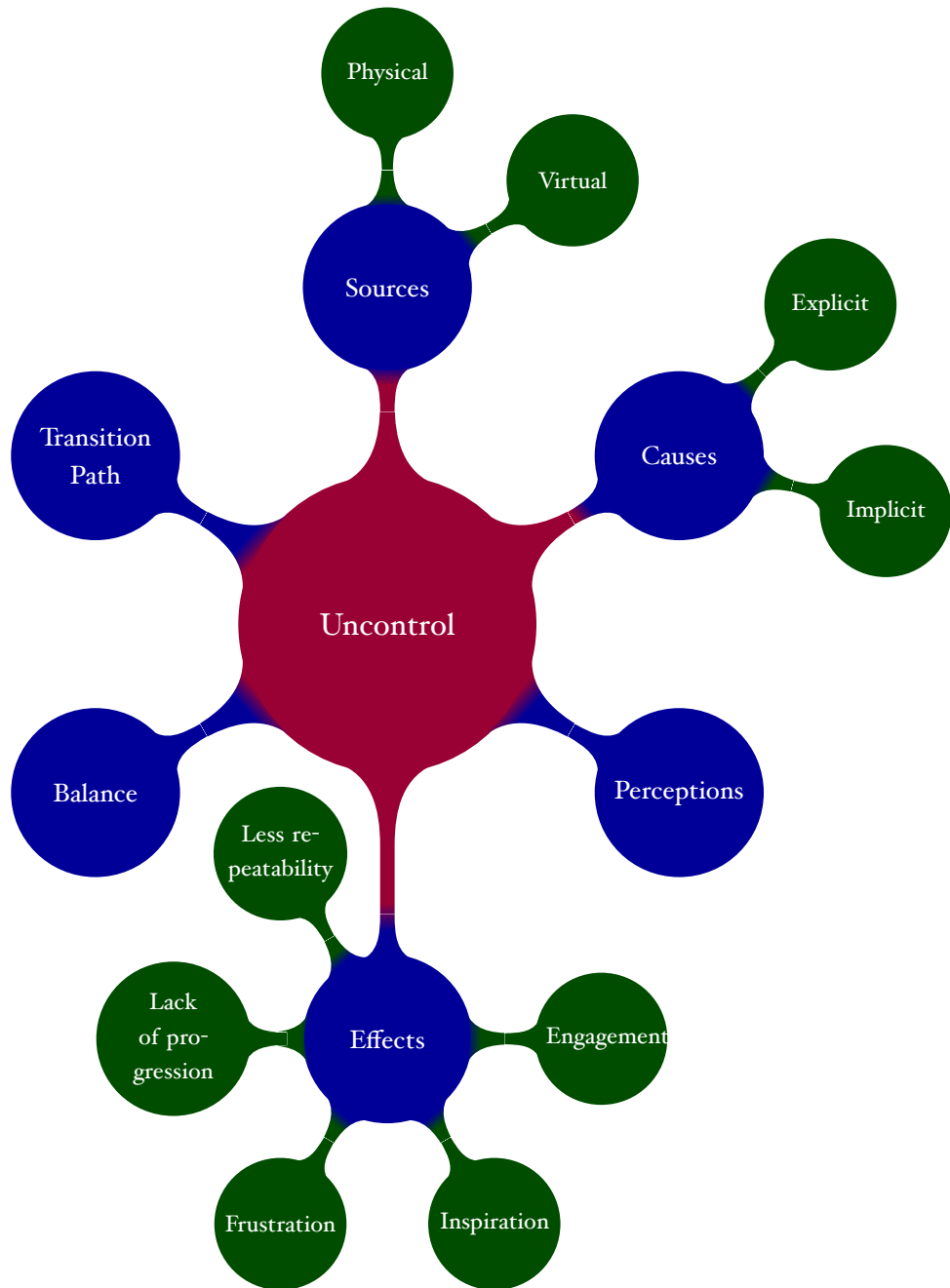


Figure 8.2: Uncontrol

is most likely because the nonlinear behaviour obscured the relationship between action and output, to an extent that it seemed to be random. This perceived randomness unfortunately deterred them from persevering with certain aspects of the system.

The effects of uncontrol are both positive and negative. On the positive side, uncontrol can lead to a heightened sense of engagement with an interface, and has the potential to model levels of embodied control as seen in acoustic instruments. In terms of the creative process, uncontrol can provide creative inspiration to the player in the form of unexpected results. On the negative side, uncontrol can lead to frustration with an interface and can interrupt or stop the learning process. It can also negatively impact repeatability.

Obtaining a good balance of uncontrol is an important factor in the design of a musical interface. The results concerning perception of uncontrol as randomness suggest that this balance is individual to each user, and is most likely related to control skill levels and also expectations of how a system will function. A level of uncontrol that is too high will not encourage the user to try and understand the dynamics of an interface.

The transition path to uncontrol describes how a system moves between control and uncontrol. This path may take many forms, from a gentle curve to a binary jump. This behaviour has a fundamental effect on the experience of using an instrument, because it determines how easily the player can control the transition from linear into nonlinear zones and back again. A gentle curve will allow the user to test out areas of uncontrol, but retreat back to safer and more linear behaviour when desired. At the opposite end of the spectrum, a binary jump into non-linear behaviour may leave the player trapped in a non-linear zone, unable to return to normal playability.

8.4.4.2 DAMPING NONLINEAR BEHAVIOUR

The evaluation results show an interesting function of RECZ, where the users employed it not just to traverse a parameter space, but also as a mechanism for damping nonlinear behaviour in mappings. Both the EchoFoam field study and the timbre space study show examples of participants narrowing the search space in order to limit or lessen unpredictability in the system. Further to this, the field study showed that this was a key element to successful use of the interface, and that even detailed control levels could be achieved as the search space was constrained. A high level search system can act as a well-matched complement to the complex behaviour of a multiparametric input device, realising a balance of controllable uncontrollability. It's reasonable to suggest that this system could act as a generic tool to enhance the behaviour of other multiparametric input devices. Furthermore, as discussed earlier, the results point to a variability in the way in which users perceive and explore elements of uncontrol. This emphasises the usefulness of the RECZ, which provides a continuous and user controllable level of damping to the system, adjustable to individual needs.

8.4.4.3 SUMMARISING UNCONTROL

The results have emphasised the significance of uncontrol in the design and use of musical interfaces. The mini-taxonomy breaks down uncontrol into smaller elements that can be analysed by the designer in order to achieve a better understanding of their system, and to

help fully exploit the benefits of uncontrol. With controllers that aim to enhance a sense of embodiment through complex motion capture, uncontrol is always going to be a key issue in many aspects of interface and mapping design. It is important that the designer provides strategies for the user to deal with positive and negative aspects of uncontrol, as it is important for the user to be aware of the significance of uncontrol elements.

8.4.5 THE CREATIVE PROCESS

As has just been discussed, both controllers exhibited a higher degree of uncontrol than is typically associated with conventional interfaces for digital music tools. Despite this, the evaluation results still showed that the majority of participants considered them to be useful controllers, albeit for a more limited range of applications. This shows that less precise controllers do have a place in the computer musician's toolbox, raising the question of how these tools best fit into the creative workflow for digital music.

There is a limited body of research available that focuses on the creative process of computer musicians. Gelineck and Serafin's (2009b) survey of electronic music composers is one of the most revealing. There are some key points in their findings that relate to the results presented here. Firstly, they mapped out the creative process and found three overall phases: an exploratory phase followed by an editing phase, and finally a pragmatic phase. The results here confirm the exploratory phase of this model, the EchoFoam field study showing that participants used the system for exploration and inspiration early on in the composition process. More significantly, Gelineck and Serafin found that the composers in their study preferred a level of unpredictability in the tools they used, and commented that designers should be aware of the levels of unpredictability in a tool and how it fits into the composition process. The results of the evaluations again confirm this, as the level of uncontrol in the interfaces had a direct impact on where the participants felt they were useful. This was especially evident in the EchoFoam study, where the majority of participants favoured the system for the explorative phase, where unpredictable behaviour would lead to interesting new results. Also, users of Phalanger felt that it was better for broad gestures rather than fine adjustment, and further to this it showed itself to be a very good device for exploration with the RECZ system.

There is one case where the results contradict Gelineck and Serafin's conclusions. They placed unpredictability on the opposite end of an axis with intuitiveness, suggesting that an unpredictable tool cannot be intuitive and vice versa. The evaluation results suggest otherwise; users from all studies reported the controllers as feeling organic, natural or intuitive in use, showing that these two concepts are not mutually exclusive. The reason for this is most likely the specialist focus of these two systems, which is on embodied interaction. There is no doubt that unpredictability in many software systems can result in behaviour which impedes intuitive use; conceivably this could happen when the user loses control of a system and cannot bring it back to normal behaviour, as a quality of the transition path between linear and non-linear zones. EchoFoam and Phalanger have been designed with the aim of exploiting innate perceptual-motor skills in interaction, putting the user in very intimate contact with the process they are controlling; this may mean that it is easier to pull back from zones of unpredictability into more linear zones. It may also mean the user has an intuitive

sense of the unpredictability in the system. This suggests that Phalanger and EchoFoam have a well defined place in the computer musician's arsenal, as tools that are aimed squarely at the exploratory phase of the composition process. Further to this, the evaluation results show that some users liked these controllers more than conventional controllers, which they felt had more limited expressive potential, or lacked an embodied connection. This shows a clear need for this class of controller, that takes a highly embodied approach to motion capture, and exhibits higher than average levels of uncontrol. The majority of input devices available for computer musicians follow the scientific instrumentation paradigm, and are more suited for editing and refinement than uncontrol and idea generation. We need more controllers that are specifically aimed towards exploration and inspiration.

Currently, with the advent of mobile music creation apps, there is a trend towards a more granular approach to music creation, using many small tools rather than following the whole composition process through a single DAW environment. This approach could work well for musical controllers as well, where an artist could reserve different input devices for different musical contexts.

To summarise, EchoFoam and Phalanger have been shown to provide an embodied and intuitive approach to less precise control. While they are not necessarily suitable for refinement and editing, there is a clear place for them in the exploration phase of the creative process, and a clear need for more interfaces that prioritise embodied experience over precision.

8.5 Summary

Five evaluation studies have been presented in this thesis. In this chapter, the results were considered together, to reflect on the evaluation methods, the design aims, and to examine the emerging themes and issues surrounding the multiparametric controllers in these studies. The next chapter presents the conclusions drawn from these results.

CHAPTER 9

CONCLUSIONS

‘I believe in deeply ordered chaos.’

Francis Bacon (Kimmelman, 1989)

9.1 Towards a Model of Multiparametric Interaction

The results in this thesis represent one of the most in-depth studies into multiparametric control since Hunt and Kirk's (2000) work. The controllers tested here are different in some significant ways: they use a higher number of parameter streams and the system does not necessarily need the user's energy for a sound to be made. However, the overall approaches are very similar, and both studies can contribute to the wider picture of multiparametric controllers. In this section, Hunt and Kirk's results will be compared to the results in this thesis, and then the results examined to see how they add to Hunt and Kirk's work. In doing this, we can establish a more detailed model of multiparametric work, to aid both designers and users.

Hunt and Kirk listed some major conclusions about their studies, and three of them are particularly relevant and comparable to the work here:

- Mappings which are not one-to-one are more engaging for users.

The results have shown repeated examples of participants having an engaged and intimate experience with the sounds they were controlling. This was particularly evident in the RECZ evaluation, where some participants felt a high sense of connectedness with the sound space they were exploring. In the evaluation of EchoFoam, participants had a distinctly different experience with the foam compared to the linear mapped sliders and in some cases felt that the foam enhanced the experience; as one participant said, *'It's more related to human touch than it's related to the very limiting one dimensional moves that you can make with the MIDI controller'*. Hunt and Kirk emphasised how their participants had fun with multiparametric control, and this has also been evident in all the evaluations here. They talked of long term potential; all the users in the EchoFoam field study said they would consider using the controller after the study finished, as did one of the Phalanger users. Generally, it's clear that multiparametric control can provide a very engaged experience, but there are also some caveats: multiparametric control is not suitable for every situation, and it's not suitable for all users.

- Complex tasks may need complex interfaces.

Hunt and Kirk's results showed that as task complexity increased, so did the performance of their multiparametric interface. There are no results here that are directly comparable as no quantitative measures have been taken, but we can look at examples in the results of similar situations, the most prominent one being the RECZ study. This gave the users the relatively complex task of navigating a ten-dimensional sound space. They did not attempt this with one-to-one mappings on a MIDI controller, so we have no evidence of how this experience would have compared. However, the results do show some participants effectively navigating this space with the two multiparametric controllers with very little training time, acknowledging a highly engaging experience. This shows a relatively complex task being performed with relative ease, so goes some way to confirming Hunt and Kirk's conclusion. Further to this, in the formative EchoFoam evaluation, some participants talked about how the foam was a more intuitive approach to the sound exploration task than the sliders, again showing the success of a multiparametric interface in a relatively complex task compared to an interface with one-to-one mappings.

- Some people prefer to think in terms of separate parameters.

It's very clear that the results here confirm this. There have been participants across the studies who expressed a preference for separable parameters. Hunt and Kirk propose

that this may be due to a preference for analytic mode thinking in some individuals. More recent research (Kim et al., 2011) suggests the existence of global interaction style preferences. This area needs to be investigated further in the context of computer music interfaces.

Overall the results here seem to confirm Hunt and Kirk's findings, and they have more to add to the overall picture of multiparametric control:

- Multiparametric control is a good approach for enhancing the quality of embodiment in musician-computer interaction.

This has already been suggested by Djajadiningrat et. al for applications in the domain of product design (Djajadiningrat et al., 2007). It can also be seen that this approach shows promise for musicians interacting with digital music tools. All studies have shown examples of participants experiencing engaged, intuitive control over musical parameters, with comments that the interfaces felt organic and natural. The two controllers exploited two natural ways of connecting with the physical environment: hand motion and manipulation of malleable material. However, these are not typically associated with the control of music, which brings us to the next point:

- The role of metaphor and sound-motion correspondence is vital in creating good mappings.

Better mappings tend to involve a perceptual connection between action and sound. With the complex level of control inherent in multiparametric interfaces, providing this connection can help the user to make sense of a complex input device and large parameter space. Conversely, mappings which make an unnatural connection between action and sound can detract from the user's experience. One way to aid the user in establishing these connections is to provide control elements that allow the freedom of motion for the user to involve imagery and metaphor in interaction. This research suggests that intangible control elements, e.g. video tracking, range sensors or accelerometers, can be used successfully for this.

- A choreographed approach to multiparametric control can enhance user experience.

When using complex body motion as input, a choreographed approach can help users create a language of interaction to interact with a complex and nonlinear parameter space. It establishes a personalised means of communication and helps to subdivide the space into manageable units.

- Designers and users should adopt a considered approach to uncontrol.

In a complex multiparametric system, a degree of uncontrol is unavoidable. Uncontrol will increase with the level of parameter codependence and nonlinearity. It's important for both designers and users to consider the level and nature of uncontrol in a system, both to take advantage of its benefits and to help manage the negative effects such as high unpredictability or reduced repeatability.

- Skill in the choice of mappings is as important as skill in physical operation of the controller.

With generic, open ended systems such as the controllers in this thesis, the choice of mappings is as important for successful operation of the controller as skill in manipulating the control streams. This is both in training a system in how to respond to motion, and in choosing the ranges and destinations of the control streams. Well considered mappings can limit complex or unpredictable behaviour.

- The multiparametric approach lends itself to the exploratory phase of music creation.

With a system that presents a complex, nonlinear parameter space to the user, an exploratory approach is typically required. This style of operation inevitably leads to unpredictable results which are of artistic interest, letting the controller be very useful for the exploratory phase of music creation. With carefully chosen mappings, these systems can also be valuable for detailed work, although this needs to be approached with a different mindset from the linear, separable controllers which are typically used for fine adjustment. With scientific style control being employed in the majority of digital music systems, multiparametric controllers can make a valuable alternative addition to the control options usually available for digital musicians.

- Meta-mapping controls can make a valuable complement to a multiparametric controller.

Systems such as RECZ or EFL software can provide a valuable set of tools to direct multiparametric control, and can help to manage some control issues. For example, RECZ can help to damp nonlinearity, and the channel muting in EFL helped to solve issues with parameter codependency. Using an interactive meta-mapping tool also encourages bimanual interaction, a natural mode of interaction for humans where one hand performs fine detailed work while the other carries out coarse adjustments (Treadaway, 2009).

- Multiparametric control does not suit all users.

Both this and Hunt and Kirk's work shows that some users prefer separable parameters. Further to this, the RECZ study and the EchoFoam field evaluation showed how some users do not seem to enjoy an explorational approach to musical control. The reasons for both these factors could be explained within the framework of cognitive styles or interaction style preferences, and more research is needed to explore this area.

- Multiparametric control lets us make intuitive sense of unintuitive parameter landscapes.

A good example of this is the control of phase modulation synthesis in the formative EchoFoam evaluation. With separable linear controllers, there is a tendency to try and match controls to perceptually meaningful synthesis parameters. The multiparametric approach gives us the opportunity to take perceptually meaningless synthesis parameters, and make intuitive sense of them through an embodied approach. This opens up the landscape for synthesis and DSP control to include obtuse and complex methods of sounds creation that we would naturally tend to avoid using.

9.2 Future Work

9.2.1 PHALANGER

Since the development of Phalanger, a key change has occurred for developers of computer vision interfaces, which is the arrival of the Microsoft's Kinect (Microsoft, 2010), an affordable 3D camera, coupled with an active development community surrounding it. The obvious progression for Phalanger is for it to be adapted to work with the Kinect as a camera. This would help to solve the key problem of lighting variance causing tracking instability, and would also enhance the possibilities of the system with rich 3D tracking.

9.2.2 ECHOFOAM

There are a number of directions this research could take in the future. Firstly, variations on the sensor could be tested; different form factors and different conductive foams. Participants in the study were asked to suggest shapes for the foam: their answers included large balls, spheres with spikes and asymmetrical shapes which would help to map locations in the foam to sounds. There were several comments on the controller's potential as a collaborative tool, and it would be intriguing to explore this as well.

Other materials, for example conductive rubber or silicone (e.g. Zheng and Shimada, 2008), exhibit similar properties to conductive foam, changing electrical resistance when deformed. The techniques outlined here could be applied to creating malleable controllers with these alternative materials, making an interesting comparison of tactile sensation.

9.2.3 A GENERIC MULTIPARAMETRIC TEMPLATE

There's a huge amount of scope to experiment with new controllers that vary the multiparametric model, but follow the same data flow, for example, groups of range sensors, continuous hall sensors, quantum tunnelling composite pressure sensors or whisker sensors (Mistree and Paradiso, 2010). All of these could be used to create interfaces which output multiple streams of data, and capture motion in a detailed way. The EchoFoam circuit board detailed in appendix A provides a generic method for reading multiple analogue voltages into a computer, and the RECZ system provides a generic way of mapping these to musical parameters.

9.2.4 MAPPING TECHNIQUES

There is still a large space to explore in the use of reservoir computing techniques, such as echo state networks, in the mapping of sensor data for human-computer interaction. ESNs are ideal for processing temporal data streams. This thesis has seen them used for dimensionality reduction. They can also be used for gesture recognition, and signal generation. Their black box operation makes them ideal for creating mapping engines that take complex sets of data streams and produce output that is useful for musical control, giving them great potential as a generic tool for dealing with complex input data.

9.2.5 EMBODIED CONTROL AND MULTIPARAMETRIC CONTROLLERS

The evaluation results highlight a number of avenues for further exploration in the area of embodied controllers and digital music tools. Language creation is of particular interest. Given a large and complex parameter space, what different cognitive strategies do musicians tend to use to create a language of interaction? How important is creativity in this language creation? And how much bearing does the nature of language creation have on specific aspects of user experience and usability? Coupled with this is the use of metaphor and imagery with controllers. This research tentatively suggests a connection between intangible control elements and the capacity for imagery, it would be interesting to explore this further. Uncontrol has been another prominent topic. The evaluation results have contributed to a more detailed picture but this is still far from complete. The path between control and uncontrol seems a

compelling and key area to explore. Is careful definition of this behaviour the best strategy for creating controllers which are highly rewarding to use? Lastly, individual differences have been a key theme in the results, with sometimes polarised reactions from participants about the controllers. What is the relationship between individual differences and preferences concerning the style of control presented to participants in this thesis? Can it be framed within cognitive style, or from other areas of research in this field? And how could this information be used? It could help individuals to choose types of controllers that suit their personal style of interaction, and also highlight areas in control style where more control options are needed.

9.3 Summary

In this chapter, the combined results from the five studies run for this thesis were reflected upon, beginning with the evaluation methodology. Two types of study were employed; formative laboratory based studies and field studies. In examining the experience of running the formative evaluations, pragmatic issues concerning aesthetic preferences, novelty factor and participant confidence were highlighted. Focusing on the field studies, key issues concerned the time taken to run the study, the open-ended nature of participant experiences, system usability and participant motivation. Overall it can be seen that these two styles of study complemented each other well; the formative evaluation obtains fast results and prepares the groundwork for the longer field study. The field study compensates for methodological shortcomings found in the formative evaluation, and can yield a rich set of data concerning the use of a controller in the artist's own workflow and environment. This style of evaluation not only helps to discover functional problems, but also paints a detailed picture of user experience which is valuable for both improving the next iteration and inspiring new designs.

Following grounded theory analysis of the interviews, a set of categories emerged describing the results. These were collected together, normalised and presented as both a summary of results and as a list of concepts concerning multiparametric interaction that may be useful to other researchers.

The EchoFoam and Phalanger controllers were presented. The evaluations provided three sets of results which were combined to build an overall picture of each controller. This was done within the context of the design framework proposed in chapter 3. The results showed that Phalanger fitted well into this framework, the main exception being fault tolerance; the system experienced difficulty in varying lighting conditions. The results also highlighted issues concerning the large degree of freedom of motions possible with this system, leading to unpredictability in use and difficulties surrounding creating a language to communicate with the system. The EchoFoam system had some significant successes in providing an intuitive mode of interaction for the exploration of parameter spaces. It also had some significant problems with accuracy and difficulty in use; participants had strong preferences as to whether they liked or disliked this malleable style of control.

Several strong themes emerged from the collected results, concerning the wider theme of multiparametric interaction. It was found that metaphor, imagery and sound-motion correspondence play a key role in the use and design of these controllers. This in turn had a bearing

on choreography and language creation, elements which should be considered by users and designers for enhancing the experience of using these devices. It was found that uncontrol was a significant factor in the use of these systems, and the results were used to build a taxonomy that highlighted the sources, causes, perceptions, effects, balances and transition path of uncontrol. They also emphasised how a considered approach should be taken to these issues by designers and users. The last emergent theme was the creative process; the results showed that there is a place for this less precise style of control in the explorative phases of musical creativity.

Finally, the results were compared with Hunt and Kirk's initial work on multiparametric control, and combined to build a more detailed picture of this style of musical interface. Multiparametric interfaces can provide a very rewarding, engaged and intimate experience of musical control for digital music tools. They also exhibit some significant difficulties in use in certain situations, but these can be reduced by user awareness of issues such as uncontrol and mapping techniques.

These controllers can provide a powerful experience in certain areas of musical control; the results point to a pluralistic approach to the control of digital music tools being very beneficial to the musician. Instead of using generic tools for every task, we can have a richer and more rewarding experience by drawing from a collection of controllers that are appropriate for different creative scenarios.

APPENDIX A

BUILDING AN ECHOFOAM CONTROLLER

A.1 Introduction

This chapter describes how to make an EchoFoam controller. The instructions describe construction in a cube shape, with 25 data streams. The design is intended to be flexible, and can be adapted as needed to make customised shapes. The circuit shown here can be expanded to output up to 64 data streams.

A.2 Materials and Tools

The Controller

- Low density ESD foam
- 32 swg enamelled copper wire
- Liquid Tape
- An Arduino board (or compatible clone)
- General purpose glue

Electronics

- Stripboard
- Electrical tape
- Heat-shrink tubing in two colours
- Single core wire, in five colours
- Two 74HC4051 analog (de)multiplexer ICs
- A 4.7M preset resistor
- Male and female PCB pin headers

Tools

- Stanley knife
- Ruler
- Soldering equipment
- Multimeter
- A disposable lighter

A.3 Instructions

A.3.1 MAKING THE CONTROLLER

1. The foam arrives in sheets. Cut this into individual squares, making enough so that they can be put together to form a cube.
2. Cut 10 equal lengths of enamelled copper wire. These will form the cable between the cube and the Arduino.

3. Scrape the enamel from the ends of the wires with a stanley knife. Test each wire with a multimeter to make sure that a contact can be made.
4. Create the first end-piece by tying five of the wires into a foam square, in the four corners and in the centre. All the wires should enter the foam through the same point, in the centre of one of the sides. After this, they can be threaded in and out of the foam until they reach the contact point, in order to create a more stable hold. At the contact point, thread the wire through and out of the foam and fold it back on itself tightly.
5. Repeat the last step to create the other end-piece.
6. Mark each foam end-piece with coloured heat-shrink tubing by placing a piece over the wires right next to the foam square. Do not melt the tubing yet as you may need to move the wires around to verify which ones belong to which contact point.
7. Make the contacts which will plug into the Arduino (see figure A.1):

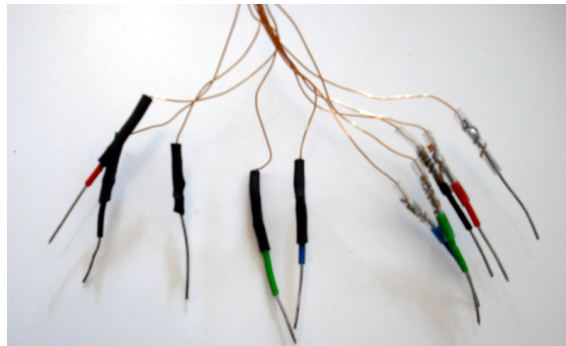


Figure A.1: Connectors

- (a) Cut 10 3 cm lengths of single core wire, two in each colour.
 - (b) Strip 1 cm from each end of the cut wires, leaving the coloured section in the centre.
 - (c) For each end-piece, take five different coloured wires. Wrap each copper wire around a single core wire and solder it on. Make sure that the colours are consistent with the contact positions of the wires for both end-pieces. Melt heat-shrink tubing around the connections, using a different colour for the contacts from each end-piece.
 - (d) Melt the heat-shrink tubing next to the foam square.
 - (e) The contacts you now have should be colour coded such that you can tell which end-piece they are attached to, and which contact area they are tied to.
8. Glue the foam squares together to create a cube, making sure the end-pieces are placed so that the contact wires are facing inwards. Place glue around the edge of each square and in the centre. Be careful not to use too much glue as this can make hard lumps in the controller.
 9. Twist the copper wires together.
 10. In a well-ventilated room, brush Liquid Tape onto the wires to form them into a single cable.

A.3.2 ELECTRONICS

This section describes how to make an Arduino shield containing the circuitry necessary to read data from the foam controller and send it to a computer.

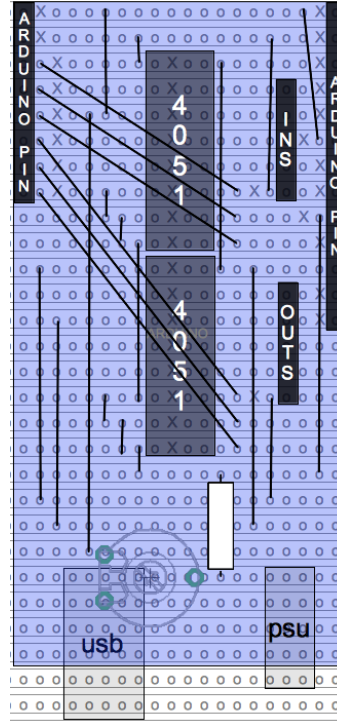


Figure A.2: EchoFoam Stripboard Layout

1. Cut out a 20 x 26 piece of stripboard.
2. Figure A.2 shows the stripboard layout, and figure A.3 shows a circuit diagram.
3. Assemble the stripboard as shown. The male pin headers should be placed through the stripboard holes so they protrude on the metal side - this way they can plug into the Arduino pin headers.
4. Place electrical tape over the USB socket on the Arduino. This insulates the socket from contact with the stripboard.
5. Plug the finished board into the Arduino (see figure A.4)
6. Plug the foam connectors into the board, one colour into the top set of pin sockets and one colour into the other set.

A.3.3 TESTING

The arduino code is listed in section B.4.9. Check the serial window to see the output from the controller, and make sure that the values change as you move the foam. Adjust the preset resistor to get a good output range; ideally values should work within the full 10 bit range of 0 - 1023.

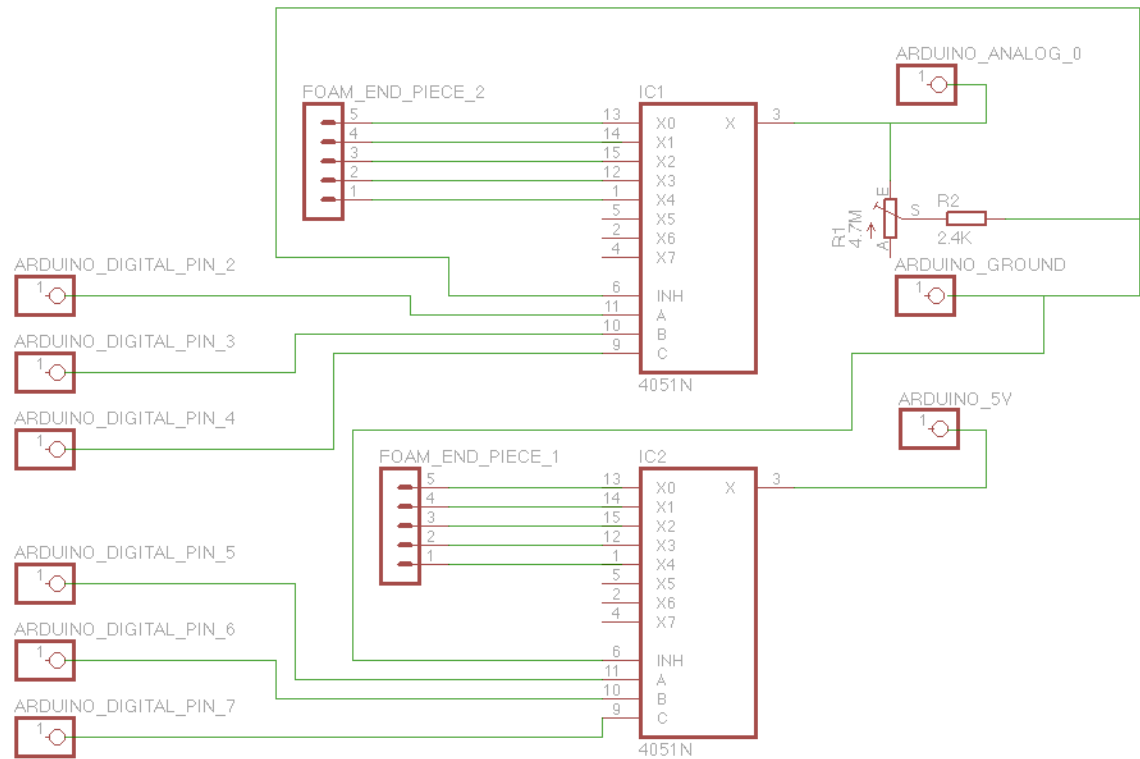


Figure A.3: EchoFoam Circuit Schematic

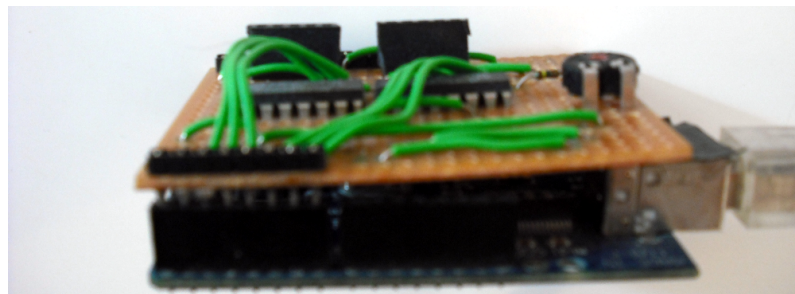


Figure A.4: The Finished Circuit Board

APPENDIX B

ECHOFOAM LIVE LINK MANUAL

This section presents the contents of a website set up for use by the participants in the EchoFoam longitudinal field evaluation. The website was created to help them to set up the controller and to provide a reference for usage instructions.

B.1 About

EchoFoam Live Link is software that lets an EchoFoam controller communicate with Ableton Live 8.x.

The EchoFoam system is a hardware/software combination for using conductive foam as a malleable controller for interaction with digital music tools. The design philosophy leans towards an exploratory approach to interacting with software, both through the nature of the controller and through mapping techniques.

The Live Link software facilitates various ways of controlling continuous parameters in Live with the EchoFoam controller.



Figure B.1: EchoFoam Live Link Screenshot

B.2 Getting Started

The first stage is to connect up the controller and get it talking to the computer, the hardware setup page has details about this.

After this, you'll need to install the EchoFoam Live Link software, and a supporting library.

B.2.1 HARDWARE SETUP

You will need to install some USB drivers to connect with the Arduino board, these can be found here: <http://www.ftdichip.com/Drivers/VCP.htm>.

To connect up the controller: there are ten wires which need to be connected to the circuit board. With the circuit board positioned so that the two sets of terminals are on the right hand side, insert the 5 wires with black wrapping into the top terminal, in the following colour order from top down: black, red, silver, green, blue. Then take the remaining wires and insert them into the lower terminal in the same colour order.

Next, plug the USB cable in and connect the controller to your computer.

B.2.2 SOFTWARE SETUP

You need libflens installed to run the software. You can download this here, then you need to copy it to somewhere in your library path such as `/usr/lib` or `/usr/local/lib`. To access these folders, type `[cmd][shift]-g` in a finder window and enter the folder path. You'll need to authorise the copy by entering your password.

The next step is to install the scripts which allow the controller to communicate with Live. Open your Live folder (probably something like `/Applications/Live 8.1.3 OS X`), ctrl-click on the `Live` application, select `Show Package Contents`, and navigate to `Contents/AppResources/MIDI Remote Scripts`. Download the scripts from here, and then unzip and copy the `LiveOSCEchoFoam` folder into this folder. This software has been developed to work with the Live 8.1.3, if you are using an older version and don't mind getting the latest update then please do. If not, the software will hopefully work fine anyway, but I can't guarantee because I haven't tested with older versions.

To configure this in Live, (re)start Live, and go to the MIDI/Sync panel on the preferences pane. Drop down one of the `Control Surface` menus, and select `LiveOSCEchoFoam` from the list.

Finally, drop the EFL folder anywhere on your harddrive, and run the `EchoFoam-LiveLink` app. To test whether the software is working, press `d` to view the incoming and outgoing data streams. Press `d` again to hide them. Any problems, please email me.

B.3 The EchoFoam Controller

The EchoFoam controller is a malleable controller; you can use it to explore parameter spaces by deforming and manipulating the controller, a different sound corresponds to each different possible deformation of the foam.

The controller is a multiparametric, non-linear and continuous streaming controller. It outputs a constant stream of 25 parallel parameters which describe the current shape of the controller. The software includes a mapping engine, which takes these 25 parameters and maps them to the number of parameters required for what you need to control in Live. These mapped parameters change continuously as the shape of the controller is manipulated, and each single position of the controller will output a consistent set of parameters.

Due to the physical nature of the controller and the way in which the parameters are mapped, the output of the controller is non-linear. Some areas of the controller will have more effect on the output than others, and you can only reach some settings by deforming one area of the foam to a certain shape and then manipulating another. All this means that you need to take an exploratory approach with this controller. The sound is embodied within the foam and you need to explore with touch to see how manipulating the foam will affect the sound. The relationship between the controller and the sound will change across different mappings, you will need to explore each one individually. This is an different to approach to the world of linear MIDI controllers, and may take some getting used to!

At more complex settings (e.g controlling ten dimensions at once), the controller can become very sensitive and possibly difficult to use. To help with this, the software provides high-level control over mapping to aid in narrowing the search space and reducing sensitivity. This same mapping logic can be use to navigate through large parameter spaces.

One practical note □ the controller is one large foam sensor and doesn't contain any internal electronics, so don't be afraid to squash it, twist it etc.

B.4 In Use

B.4.1 CAUTIONARY ADVICE

This software is a beta, and realistically there will be some usage scenarios I haven't covered in development, and so possibly some bugs. To this extent, please use this on a copy of your Live sets, rather than the original. Just in case □

B.4.2 SCOPE / LIMITATIONS

Live Link will let you modulate continuous parameters on devices within Ableton Live. Mappings unfortunately will not work with return or master tracks, but will work with normal tracks.

B.4.3 OVERVIEW

The basic function of Live Link is to connect an EchoFoam controller to Live, to control continuous parameters. Expanding on this basic function, there are different ways of choosing parameters to control, and options for controlling how the controller maps onto these parameters.

Data arrives at the software from the controller as 25 continuous streams. This raw data needs refining before it can be used, so it's passed through a mapping process (using a recurrent

Echo State Network), which outputs the number of streams which you require for controlling parameters in Live. You can choose to map to between one and ten streams. These data streams are non-linear and interdependent; only through exploring the mappings by manipulating the controller will you be able to determine exactly what the mappings do.

B.4.4 USER INTERFACE

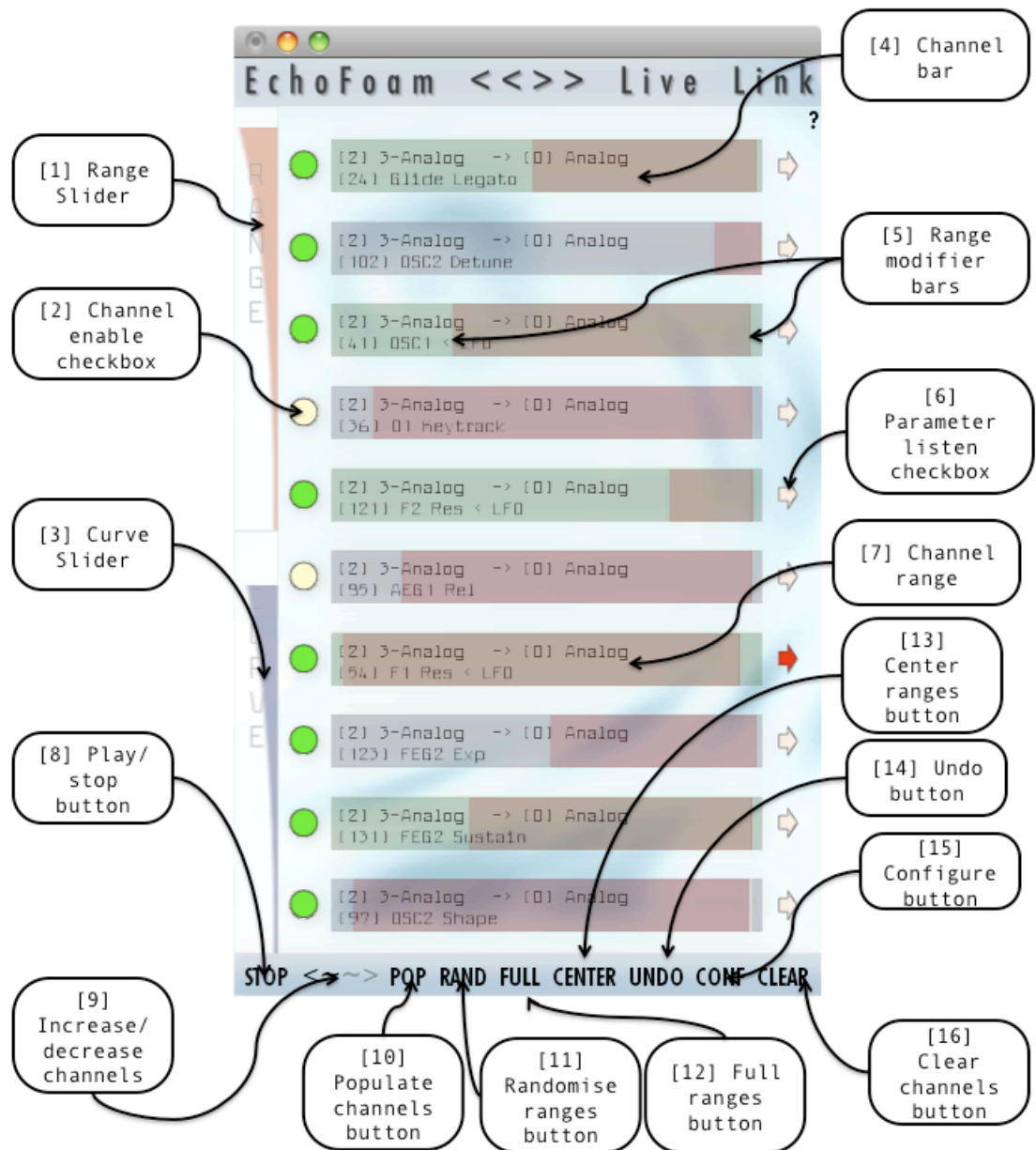


Figure B.2: EchoFoam Live Link User Interface

1. Range Slider

This slider modifies the width of the channel ranges, changing all ranges on all channels by the same factor.

2. Channel enable checkbox

When this is green, data is being sent to Live; when it's yellow, no data is sent.

3. Curve Slider

This changes the chance that random ranges will be large. Set high to get (mostly) large random ranges, and low to get (mostly) small ranges.

4. Channel bar

This represents a channel, where data is received from the mapping process and then sent to Live; it shows the parameters that the channel is mapped to and the working range of the channel. You can click on a channel to select it; a selected channel is indicated by a yellow bar on the left hand edge. When one or more channels are selected, the populate, random, centre and clear buttons will only apply to the selected channels.

5. Range modifier bars

Drag these to manually change ranges.

6. Parameter listen checkbox

Click to arm a channel; Live Link will listen for the next device parameter to be modified in Live and set this as the channel's parameter.

If pressed by mistake, click again to disarm.

7. Channel range

This represents the working range for this channel.

8. Play/stop button

Click to stop and start data being sent to Live.

9. Increase/decrease channel buttons

Click to increase and or decrease the number of channels in use.

10. Populate channels button

Click to randomly populate the channels with parameters from the currently selected device in Live.

11. Randomise ranges button

Randomise the channel working ranges.

12. Full ranges button

Click to maximise ranges on all channels.

13. Centre ranges button

This button centres the working ranges on the current value in each channel.

14. Undo button

Undoes the last action on the working ranges.

15. Configure button

Shows the configuration panel.

16. Clear channels button

Clears all the channels.

B.4.5 CHOOSING PARAMETERS

There are two ways of choosing parameters to control

- Choosing single parameters
- Populating with randomly chosen parameters

Any device which you choose to control in Ableton will be tagged with a number in its title which looks something like ((238974234)). This is so Live Link can keep track of the devices it connects to when tracks and devices move around within the song. Please do not delete these tags, you can edit the rest of the title though as long as you do not edit the tag.

B.4.5.1 CHOOSING SINGLE PARAMETERS

To select a single parameter, arm a channel by clicking on the arrow button to the right, and then go to Live and move the parameter you'd like to assign. The details of this parameter should now appear in the channel's display.

B.4.5.2 POPULATING WITH RANDOMLY CHOSEN PARAMETERS

The 'Pop' (Populate) button on the bottom button bar selects a set of random parameters from the currently selected device in Live. This facilitates random exploration of the parameter space of a device, which can yield both pleasantly and unpleasantly surprising results! Live Link knows nothing about the meaning of parameters in the device you are selecting them from; it may start to modulate master volumes or module on/off switches, resulting in little or no sound. On the other hand it may select a serendipitous set of parameters, and a really great soundspace to explore! Use at your own risk ☐

Note that there are some issues with the Live API with detecting the currently selected device. The 'Pop' button will appear greyed out if it doesn't know which is the current device. Sometimes changing to another track and then back again will help Live Link to detect the device.

B.4.6 CONTROLLING THE SEARCH SPACE

Live Link provides high level control of the parameter space in which the controller is working, which can be used for navigating through a large search space, and narrowing it down to find a sound you want to use.

Firstly it's necessary to introduce the concept of the working range. This is the sub-range of a parameter's total range within which one output from the controller can move, for example between 15% and 67%. Live Link lets you manipulate the working range of each channel, thereby letting the controller work within constrained sections of a larger multi-dimensional parameter space. Each set of working ranges, depending on the sound the channels are mapped to, will give a different feeling to the manipulation of the controller.

The simplest way to modify the working range is to hit the Rand button, which randomises all the working ranges. Hitting Full maximises all the working ranges. The curve slider changes

the chance of a random range being large; to explore larger areas of a parameter space move this slider towards the top. to explore narrow sections, lower the value of a slider.

The process of manipulating ranges can be used to navigate a search space, zooming in on the sound you want. When you have the controller moved such that the kind of sound you are aiming for is being heard, hit the centre button and this will centre all the working ranges on this point. Moving the range slider will narrow these ranges, limiting the working area to sounds closer to the centre of the ranges. Repeating the pattern of explore-centre-zoom lets you navigate and narrow the search space until you reach the desired sound (or press Rand to start over).

B.4.7 KEY COMMANDS

'd': debug mode. This shows two data graphs; the first is the 25 sensor readings from the EchoFoam controller, and the second is the outputs from the mapping process.

'r': rand
 [return]: center
 [left arrow]: undo
 'f': full range
 'p': populate
 '1': toggle selection for channel 1
 '2': toggle selection for channel 2
 '3': toggle selection for channel 3
 '4': toggle selection for channel 4
 '5': toggle selection for channel 5
 '6': toggle selection for channel 6
 '7': toggle selection for channel 7
 '8': toggle selection for channel 8
 '9': toggle selection for channel 9
 'o': toggle selection for channel 10
 '-': invert channel selection
 [backspace]: clear channel selection

B.4.8 ADJUSTING CONTROLLER SENSITIVITY

The foam from which the controller is made can sometimes change over time, particularly with heavy use, and start to lose sensitivity. To compensate for this, the sensitivity can be increased in two ways.

1. The configuration panel. You can move the gain slider to change the sensitivity.
2. The sensitivity dial on the circuit board. If your circuit board has a small potentiometer near the USB port, you can move the central dial with a screw driver to adjust the sensitivity.

When adjusting the sensitivity, hit the □d' key to show the input from the sensor (in the top half of the screen), and adjust so that the parameter streams take up roughly the middle section of the graph.

B.4.9 USING THE ARDUINO FOR OTHER PROJECTS

The Arduino can be used for other projects; this will require uploading new code to the board so if you want to change it back to work with the EchoFoam controller, you'll need this script.

```
int val;
int addr[] [3] = {{0,0,0},{1,0,0},{0,1,0},{1,1,0},{0,0,1},{1,0,1},{0,1,1},{1,1,1}};

void setup() {
    pinMode(2, OUTPUT);
    pinMode(3, OUTPUT);
    pinMode(4, OUTPUT);
    pinMode(5, OUTPUT);
    pinMode(6, OUTPUT);
    pinMode(7, OUTPUT);
    Serial.begin(115200);
}

void loop() {
    for(int j=0; j <= 4; j++) {
        digitalWrite(5, addr[j][0]); //1
        digitalWrite(6, addr[j][1]); //2
        digitalWrite(7, addr[j][2]); //4
        for(int i=0; i <= 4; i++) {
            digitalWrite(2, addr[i][0]); //1
            digitalWrite(3, addr[i][1]); //2
            digitalWrite(4, addr[i][2]); //4
            delay(0.000655);
            val = analogRead(0);
            Serial.print(val, DEC);
            Serial.print(" ");
        }
        Serial.println("");
        delay(10);
    }
}
```

The EchoFoam circuit board is an 'Arduino Shield' which can simply be pulled away from the Arduino and put back when required.

APPENDIX C

COLLECTED CONCEPTS FROM INTERVIEW ANALYSES

C.1 Collected Concepts from all Interview Results

Abstraction

- Affordances
- Applications
- Attention
- Bimanual Interaction
- Body
- Calibration
- Channels
- Cognitive Load
- Comparative Interfaces
- Comparison
- Composition
- Constraints
- Continuous Control
- Control
- Controllability
- Controller Design
- Controller Exploration
- Controller Physicality
- Creative Process
- Creativity
- Data Flow
- Degrees of Freedom
- Difficulty
- Dimensionality
- Embodiment
- Engagement
- Enjoyment
- Environment
- Ergonomics
- Exploration
- Expressiveness
- Familiarity
- Feature Request
- Feedback
- Fragility
- Fun
- Gesture Recognition
- Hand
- High Dimensional Control
- Imprecision
- Instrumentness
- Interaction Style
- Intuitiveness
- Language

Latency
Learnability
Learning Curve
Mapping
Metaphor
Methodology
Motion
Motions
Multiparametric Control
Musical Practice
Musicality
Negativity
Novelty
Performance
Physical
Physical Feedback
Physical Manipulation
Positivity
Precision
Randomisation
Ranges
Reliability
Repeatability
Responsiveness
Usability
Usage
Software-Hardware Interactions
Sound Exploration
Sound-Motion Correspondence
Suggestion
Tangibility
Training
Triggering
Trust
Usage
Visual Feedback
Visual Reference
Workflow

BIBLIOGRAPHY

- Ankur Agarwal, Shahram Izadi, Manmohan Chandraker, and Andrew Blake. High precision multi-touch sensing on surfaces using overhead cameras. In *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer System*, 2007.
- Michael L. Anderson. Embodied cognition: A field guide. *Artificial Intelligence*, 149:91–130, 2003.
- Alissa Antle, Greg Corness, and Milena Droumeva. Human-computer-intuition? exploring the cognitive basis for intuition in embodied interaction. *International Journal of Arts and Technology*, 2(3), 2009.
- Daniel Arfib, Jean-Michel Couturier, and Loïc Kessous. Expressiveness and digital musical instrument design. *Journal of New Music Research*, 2005.
- Newton Armstrong. *An Enactive Approach to Digital Musical Instrument Design*. PhD thesis, Princeton University, 2006.
- Daniel Ashbrook and Thad Starner. Magic: a motion gesture design tool. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, pages 2159–2168, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-929-9. doi: <http://doi.acm.org/10.1145/1753326.1753653>.
- Massimo Banzi. *Getting Started with Arduino*. Make, 2009.
- Randall D. Beer. On the dynamics of small continuous-time recurrent neural networks. *Adaptive Behavior*, 3(4):469–509, 1995.
- Todd Bentley, Lorraine Johnston, and Karola von Baggo. Evaluation using cued-recall debrief to elicit information about a user's affective experiences. In *OZCHI '05: Proceedings of the 17th Australia conference on Computer-Human Interaction*, pages 1–10, Narrabundah, Australia, Australia, 2005. Computer-Human Interaction Special Interest Group (CHISIG) of Australia. ISBN 1-59593-222-4.
- Joshua Berger and Eric Lengvenis. An interview with trent reznor. <http://www.theninhotline.net/archives/articles/plazm94a.shtml>, 1994.
- Olav W. Bertelsen, Morten Breinbjerg, and Søren Pold. Instrumentness for creativity mediation, materiality & metonymy. In *C&C '07: Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*, pages 233–242, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-712-4. doi: <http://doi.acm.org/10.1145/1254960.1254992>.
- Olav W. Bertelsen, Morten Breinbjerg, and Søren Pold. Emerging materiality: Reflections on creative use of software in electronic music composition. *Leonardo*, 42(3):97–202, 2009.
- C. Bishop. *Neural networks for pattern recognition*. Oxford: Clarendon Press, 1995.
- Alan Blackwell and Nick Collins. The programming language as a musical instrument. In *In Proceedings of PPIGo5 (Psychology of Programming Interest Group)*, pages 120–130, 2005.
- Oliver Bown, Alice Eldridge, and Job McCormack. Understanding interaction in contemporary digital music: from instruments to behavioural objects. *Organised Sound*, 14(2):188–196, 2009.
- John Seely Brown, Allan Collins, and Paul Duguid. Situated cognition and the culture of learning. *Educational Researcher*, 1989.

- Bernd Bruegge, Christoph Teschner, Peter Lachenmaier, Eva Fenzl, Dominik Schmidt, and Simon Bierbaum. Pinocchio: conducting a virtual symphony orchestra. In *ACE '07*, pages 294–295, New York, NY, USA, 2007. ACM Press. ISBN 978-1-59593-640-0. doi: <http://doi.acm.org/10.1145/1255047.1255132>.
- Nick Bryan-Kinns and Fraser Hamilton. Identifying mutual engagement. *Behaviour and Information Technology*, 31(2):101–125, 2009. doi: [10.1080/01449290903377103](https://doi.org/10.1080/01449290903377103). URL <http://www.tandfonline.com/doi/abs/10.1080/01449290903377103>.
- Hans-Jorg Bullinger, Jurgen Ziegler, and Wilhelm Bauer. Intuitive human-computer interaction—toward a user-friendly information society. *International Journal of Human-Computer Interaction*, 14(1):1–23, 2002. ISSN 10447318.
- Sebastian Burkhart and Tobias Moebert. The wiinstrument. <http://sourceforge.net/projects/wiinstrument/>, 2007.
- Anne-Marie Burns and Marcelo M. Wanderley. Visual methods for the retrieval of guitarist fingering. In *NIME '06: Proceedings of the 2006 conference on New interfaces for musical expression*, pages 196–199, Paris, France, France, 2006. IRCAM — Centre Pompidou. ISBN 2-84426-314-3.
- Andreas Burzik. Go with the flow. *The Strad*, 2003.
- Richard Buskin. Classic tracks: Orbital 'chime'. *Sound On Sound*, December 2006.
- Eliot Van Buskirk. Controlling a synthesizer with the wii remote. http://www.wired.com/listening_post/2006/12/controlling_a_s/, 2006.
- John Butcher, David Verstraeten, Benjamin Schrauwen, Charles Day, and Peter Haycock. Extending reservoir computing with random static projections: a hybrid between extreme learning and rc. In *European Symposium on Artificial Neural Networks*, 2010.
- W. Buxton. Artists and the art of the luthier. *Computer Graphics: The SIGGRAPH Quarterly*, 31(1):10–11, 1997.
- Antonio Camurri, Paolo Coletta, Giovanna Varni, and Simone Ghisio. Developing multimodal interactive systems with eyesweb xmi. In *NIME '07: Proceedings of the 7th international conference on New interfaces for musical expression*, pages 305–308, New York, NY, USA, 2007. ACM. doi: <http://doi.acm.org/10.1145/1279740.1279806>.
- Sergio Cavaliere, Gianfranco Evangelista, and Aldo Piccialli. Synthesis by phase modulation and its implementation in hardware. *Comput. Music J.*, 12:29–42, March 1988. ISSN 0148-9267. doi: <http://dx.doi.org/10.2307/3679835>. URL <http://dx.doi.org/10.2307/3679835>.
- Joel Chadabe. Interactive composing: An overview. *Computer Music Journal*, 1984.
- Angela Chang and Hiroshi Ishii. Zstretch: a stretchy fabric music controller. In *NIME '07: Proceedings of the 7th international conference on New interfaces for musical expression*, pages 46–49, New York, NY, USA, 2007. ACM. doi: <http://doi.acm.org/10.1145/1279740.1279746>.
- Chih-Chung Chang and Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Noel Chateau and Marc Mersiol. Amuse: A tool for evaluating affective interfaces. In *CHI'05 Workshop on Evaluating Affective Interfaces - Innovative Approaches*, 2005. URL http://www.sics.se/~kia/evaluating_affective_interfaces/Chateau.pdf.
- H. Chen, R.T. Wigand, and M.S. Nilan. Optimal experience of web activities. *Computers in Human Behavior*, 1999.
- Jonny Lee Chung. Hacking the nintendo wii remote. *IEEE Pervasive Computing*, 2008.
- Clark and Chalmers. The extended mind. *The Philosopher's Annual*, vol XXI, 1998. URL <http://consc.net/papers/extended.html>.

- Andy Clark. *Being There: Putting Brain, Body and World Together Again*. MIT Press, 4th edition, 1997.
- N. Collins, A. McLean, J. Rohrerhuber, and A. Ward. Live coding in laptop performance. *Organised Sound*, 2003.
- Nicolas Collins. Exploded view: the musical instrument at twilight. <http://www.steim.org/steim/texts.php?id=23>, 1993.
- Nicolas Collins. *Handmade Electronic Music: The Art of Hardware Hacking*. Routledge, 2nd edition, 2009.
- Arshia Cont, Thierry Coduys, and Cyrille Henry. Real-time gesture mapping in pd environment using neural networks. In *NIME '04: Proceedings of the 2004 conference on New Interfaces for Musical Expression*, 2004.
- Perry Cook. Principles for designing computer music controllers. In *NIME '01: Proceedings of the 2001 conference on New interfaces for musical expression*, pages 1–4, Singapore, Singapore, 2001. National University of Singapore.
- Tim Coughlan and Peter Johnson. Interaction in creative tasks. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 531–540, New York, NY, USA, 2006. ACM. ISBN 1-59593-372-7. doi: <http://doi.acm.org/10.1145/1124772.1124854>.
- Mihály Csíkszentmihályi. *Flow: The Psychology of Optimal Experience*. HarperPerennial, 1991.
- Palle Dahlstedt. Evolution in creative sound design. In *Evolutionary Computer Music*. Springer London, 2007.
- Palle Dahlstedt. Dynamic mapping strategies for expressive synthesis performance and improvisation. In *Computer Music Modeling and Retrieval. Genesis of Meaning in Sound and Music: 5th International Symposium, CMMR 2008 Copenhagen, Denmark, May 19-23, 2008 Revised Papers*, pages 227–242. Springer-Verlag, Berlin, Heidelberg, 2009. ISBN 978-3-642-02517-4. doi: http://dx.doi.org/10.1007/978-3-642-02518-1_16.
- Örjan de Manzano, Töres Theorell, László Harmat, and Fredrik Ullén. The psychophysiology of flow during piano playing. *Emotion*, 10(3):301–311, June 2010.
- Gal Détourn. Interview: Orbital. *Making Music*, (181), May 2001.
- John K. Diliberto. An interview with donald buchla. *Polyphony*, pages 14–17, 1983.
- Alan Dix, Janet Finlay, Gregory D. Abowd, and Russel Beale. *Human-Computer Interaction*. Prentice Hall, 3rd edition, 2004.
- DJ WüJ. DJ WüJ. <http://www.djwüj.com/>, 2010.
- Tom Djajadiningrat, Ben Matthews, and Marcelle Stienstra. Easy doesn't do it: skill and expression in tangible aesthetics. *Personal Ubiquitous Comput.*, 11(8):657–676, 2007. ISSN 1617-4909. doi: <http://dx.doi.org/10.1007/s00779-006-0137-9>.
- Paul Dourish. *Where The Action Is: The Foundations of Embodied Interaction*. The MIT Press, 1st edition, 2001.
- Barry Eaglestone, Nigel Ford, Peter Holdridge, and Jenny Carter. Are cognitive styles an important factor in design of electroacoustic music software?†. *Journal of New Music Research*, 37(1):77 – 85, 2008. URL <http://www.informaworld.com/10.1080/09298210802259561%0929-8215%20September19,2008>.
- Brian Eno. The studio as compositional tool. In *Audio Culture: Readings In Modern Music*. Continuum Books, 2002.
- Brian Eno and Peter Schmidt. *Oblique Strategies (Over One Hundred Worthwhile Dilemmas)*. Opal OBLIQUE CARDS, Opal Records, 1975, 1978.

- Stephen H. Fairclough. Fundamentals of physiological computing. *Interacting with Computers*, 2008. ISSN 0953-5438. doi: DOI:10.1016/j.intcom.2008.10.011. URL <http://www.sciencedirect.com/science/article/B6V0D-4TW6HNY-1/2/985ac0391a0fd183dcf6a7886db8de21>.
- Daniel Fallman and John Waterworth. Dealing with user experience and affective evaluation in hci design: A repertory grid approach. In *CHI'05 Workshop on Evaluating Affective Interfaces - Innovative Approaches*, 2005.
- Chrisantha Fernando and Sampsa Sojakka. Pattern recognition in a bucket. *Advances in Artificial Life*, pages 588-597, 2003.
- Rebecca Fiebrink. *Real-time Human Interaction with Supervised Learning Algorithms for Music Composition and Performance*. PhD thesis, Princeton University, 2011.
- Nigel Ford, T.D. Wilson, Allen Foster, David Ellis, and Amanda Spink. Information seeking and mediated searching. part 4. cognitive styles in information seeking. *Journals for the American Society for Information Science and Technology*, 53(9):728-735, 2002.
- David Forsyth and Jean Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, 2003.
- Adrian Freed. Application of new Fiber and Malleable Materials for Agile Development of Augmented Instruments and Controllers. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 107-112, 2008.
- Steve Gelineck and Stefania Serafin. A quantitative evaluation of the differences between knobs and sliders. In *NIME '09: Proceedings of the 9th international conference on New interfaces for musical expression*, 2009a.
- Steven Gelineck and Stefania Serafin. From idea to realization - understanding the compositional processes of electronic musicians. In *Audio Mostly*, 2009b.
- Reed Ghazala. *Circuitbending: Build Your Own Alien Instruments*. John Wiley & Sons, 2005.
- Satinder P. Gill. Entrainment and musicality in the human system interface. *AI and Society*, 21(4): 567-605, 2007.
- Barney Glasser and Anselm Strauss. *Discovery of Grounded Theory. Strategies for Qualitative Research*. Sociology Press, 1967.
- Albert Glinsky. *Theremin: Ether Music and Espionage*. University of Illinois Press, 2005.
- J.P. Guilford. Cognitive styles: What are they? *Educational and Psychological Measurement*, 40(3):715-735, 1980.
- Kjetil F. Hansen and Marcos Alonso. More DJ Techniques on the reactable. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2008.
- Richard L. Hazlett. Using biometric measurement to help develop emotionally compelling games. In *Game Usability*. Focal Press, 2007.
- Helen Heath and Sarah Cowley. Developing a grounded theory approach: a comparison of glaser and strauss. *International Journal of Nursing Studies*, 41(2):141 - 150, 2004. ISSN 0020-7489. doi: DOI:10.1016/S0020-7489(03)00113-5. URL <http://www.sciencedirect.com/science/article/B6T7T-49FGKCG-1/2/1cbe16a75b12b77b1f13224e1b97db35>.
- Michael Hohl. Designing the art experience: using grounded theory to develop a model of participants' perception of an immersive telematic artwork. *Digital Creativity*, 20(3):187-196, September 2009.
- Georg Holzmann. Reservoir computing: a powerful black-box framework for nonlinear audio processing. In *DAFx*, 2009.

- Jonathan Hook, Stuart Taylor, Alex Butler, Nicolas Villar, and Shahram Izadi. A reconfigurable ferromagnetic input device. In *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 51–54, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-745-5. doi: <http://doi.acm.org/10.1145/1622176.1622186>.
- Kristina Höök, Phoebe Sengers, and Gerd Andersson. Sense and sensibility: evaluation and interactive art. In *CHI '03*, pages 241–248, New York, NY, USA, 2003. ACM. ISBN 1-58113-630-7. doi: <http://doi.acm.org/10.1145/642611.642654>.
- Ming-Kuei Hu. Visual pattern recognition by moment invariants. *IRE Transactions On Information Theory*, 1962.
- Andy Hunt and Ross Kirk. Mapping strategies for musical performance. In M.M. Wanderley and M. Battier, editors, *Trends in Gestural Control of Music*. Ircam - Centre Pompidou, 2000.
- Andy Hunt, Marcelo M. Wanderley, and Matthew Paradis. The importance of parameter mapping in electronic instrument design. In *Proceedings of the Conference on New Instruments for Musical Expression*, 2002.
- Tom Igoe and Dan O'Sullivan. *Physical Computing: Sensing and Controlling the Physical World with Computers*. Course Technology PTR, 2004.
- Don Ihde. Technologies–musics–embodiments. *Janus Head*, 2007. URL <http://www.janushead.org/10-1/Ihde.pdf>.
- Hiroshi Ishii and Brygg Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, 1997.
- K. Ishu, T. van der Zant, V. Becanovic, and P. Ploger. Identification of motion with echo state network. In *IEEE Techno-Ocean*, volume 3, pages 1205 – 1210 Vol.3, nov. 2004. doi: 10.1109/OCEANS.2004.1405751.
- Julie Jacko and Andrew Sears. *The Human-Computer Interaction Handbook*. LEA, 2003.
- JazzMutant. Jazzmutant: Richard devine. http://www.jazzmutant.com/artists_devine.php, August 2010.
- Alexander Refsum Jensenius. *Action - Sound: Developing Methods and Tools to Study Music-Related Body Movement*. PhD thesis, University of Oslo, 2007. URL <http://folk.uio.no/alexanje/phd/jensenius-phd.pdf>.
- Paula Skutnick-Henley Joann Marie Kirchner, Arvid J. Bloom. The relationship between performance anxiety and flow. *Medical Problems of Performing Artists*, 23(2):59, 2008.
- B. Jones, D. Stekel, J. Rowe, and C. Fernando. Is there a liquid state machine in the bacterium *escherichia coli*? In *Artificial Life*, 2007.
- Sergi Jordà. *Digital Lutherie: Crafting musical computers for new musics' performance and improvisation*. PhD thesis, Universitat Pompeu Fabra, 2005.
- Sergi Jordà. On stage: the reactable and other musical tangibles go real. *International Journal of Arts and Technology*, 2009.
- Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. The reactable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Tangible and embedded interaction*, 2007. ISBN 978-1-59593-619-6. doi: <http://doi.acm.org/10.1145/1226969.1226998>.
- Joseph 'Jofish' Kaye, Kirsten Boehner, Jarmo Laaksolahti, and Anna Staahl. Evaluating experience-focused hci. In *CHI '07: CHI '07 extended abstracts on Human factors in computing systems*, pages 2117–2120, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-642-4. doi: <http://doi.acm.org/10.1145/1240866.1240962>.

- David Keane. Some practical aesthetic problems of electronic music composition. *Journal of New Music Research*, 21(1):9 – 19, 1979.
- Judy Kendall. Axial coding and the grounded theory controversy. *Western Journal of Nursing Research*, 21:743–757, 1999.
- Chris Kiefer. Input devices and mapping techniques for the intuitive control of composition and editing for digital music. In *TEI '10: Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, pages 311–312, New York, NY, USA, 2010a. ACM. ISBN 978-1-60558-841-4. doi: <http://doi.acm.org/10.1145/1709886.1709958>.
- Chris Kiefer. Exploring timbre spaces with two multiparametric controllers. In *Proceedings of Sound and Music Computing*, 2010b.
- Chris Kiefer. A malleable interface for sonic exploration. In *Proceedings of New Interfaces for Musical Expression*, 2010c.
- Chris Kiefer, Nick Collins, and Geraldine Fitzpatrick. HCI methodology for evaluating musical controllers: A case study. In *Proceedings of New Interfaces for Musical Expression*, 2008a.
- Chris Kiefer, Nick Collins, and Geraldine Fitzpatrick. Evaluating the wiimote as a musical controller. In *Proceedings of the International Computer Music Conference*, 2008b.
- Chris Kiefer, Nick Collins, and Geraldine Fitzpatrick. Phalanger: Controlling music software with hand movement using a computer vision and machine learning approach. In *New Interfaces For Musical Expression*, 2009.
- Da-jung Kim, Youn-kyung Lim, and Hyeon-Jeong Suk. My own-style interaction: exploring individuals' preferences to interactivity. In *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*, CHI EA '11, pages 1861–1866, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0268-5. doi: <http://doi.acm.org/10.1145/1979742.1979897>.
- Michael Kimmelman. Unnerving art: Francis bacon at 80. *The New York Times Magazine*, August 1989.
- Peter Kirn. Playing live, dangerously: Mouse on mars as augmented band. <http://createdigitalmusic.com/2012/06/playing-live-dangerously-mouse-on-mars-as-augmented-band-video/>, June 2012.
- David Kirsh and Paul Maglio. On distinguishing epistemic from pragmatic action. *Cognitive Science*, 1994.
- J. D. Landis. *Longing*. Snowbooks, 2000.
- Victor Lazzarini, Joseph Timoney, and Thomas Lysaght. The generation of natural-synthetic spectra by means of adaptive frequency modulation. *Computer Music Journal*, 32(2), 2008.
- S.J. Lederman and R.L. Klatzky. Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19(3):342–368, July 1987.
- Iara Lee. Interview: Teo macero. <http://www.furious.com/perfect/teomacero.html>, 1997.
- Xiaowei Lin, Zehong Yang, and Yixu Song. The application of echo state network in stock data mining. In Takashi Washio, Einoshin Suzuki, Kai Ting, and Akihiro Inokuchi, editors, *Advances in Knowledge Discovery and Data Mining*, volume 5012 of *Lecture Notes in Computer Science*, pages 932–937. Springer Berlin / Heidelberg, 2008. URL http://dx.doi.org/10.1007/978-3-540-68125-0_95.
- Seth Lloyd. Computational capacity of the universe. *Phys. Rev. Lett.*, 88(23):237901, May 2002. doi: 10.1103/PhysRevLett.88.237901.
- Pedro Lopes, Alfredo Ferreira, and J. A. Madeiras Pereira. Battle of the djs: an hci perspective of traditional, virtual, hybrid and multitouch djing. In *NIME '12: Proceedings of the 12th international conference on New interfaces for musical expression*, 2012.

- Mantas Lukosevicius and Herbert Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 2009.
- Ian MacDonald. *River of the Gods*. Gollancz, 2004.
- Raymond MacDonald, Charles Byrne, and Lana Carlton. Creativity and flow in musical composition: an empirical investigation. *Psychology of Music*, 34(3):292–306, 2006. doi: 10.1177/0305735606064838. URL <http://pom.sagepub.com/cgi/content/abstract/34/3/292>.
- Mary-Anne Mace and Tony Ward. Modeling the creative process: A grounded theory analysis of creativity in the domain of art making. *Creativity Research Journal*, 14(2):179–192, April 2002.
- Thor Magnusson. *Epistemic Tools: The Phenomenology of Musical Instruments*. PhD thesis, University of Sussex, 2009.
- Thor Magnusson and Enrike Hurtado Mendieta. The acoustic, the digital and the body: a survey on musical instruments. In *NIME '07: Proceedings of the 7th international conference on New interfaces for musical expression*, pages 94–99, New York, NY, USA, 2007. ACM. doi: <http://doi.acm.org/10.1145/1279740.1279757>.
- Regan Mandryk. Physiological measures for game evaluation. In *Game Usability: Advancing the Player Experience*. Morgan Kaufmann, 2008.
- Regan L. Mandryk, Kori M. Inkpen, and Thomas W. Calvert. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour and Information Technology*, 2006.
- Regan Lee Mandryk. Evaluating affective computing environments using physiological measures. In *CHI'05 Workshop on Evaluating Affective Interfaces - Innovative Approaches*, 2005.
- Martin Marier. The sponge: A flexible interface. In *NIME '10: Proceedings of the 10th international conference on New interfaces for musical expression*, 2010.
- James McCartney. Rethinking the computer music language: SuperCollider. *Computer Music Journal*, 26(4):61–8, 2002.
- Josh McDermott. The evolution of music. *Nature*, 453:287–288, May 2008.
- Microsoft. Microsoft kinect. <http://www.xbox.com/kinect>, 2010.
- Matthias Milczynski, Thomas Hermann, Till Bovermann, and Helge Ritter. A malleable device with applications to sonification-based data exploration. In *Proceedings of the 12th International Conference on Auditory Display*, 2006.
- Greg Milner. *Perfecting Sound Forever*. Granta Publications, 2009.
- Pejman Mirza-Babaei and Graham McAllister. Using biometrics to better understand the players' gameplay experience. In *Proceedings of Future Game On Conference*, Parc des Expositions, Paris, pavilion 7., September 9–10 2010a.
- Pejman Mirza-Babaei and Graham McAllister. Using physiological measures in conjunction with other usability approaches for better understanding of the player's gameplay experiences. In *Game Research Methods Seminar*, Tampere, Finland, April 8–9 2010b.
- Behram F.T. Mistree and Joseph A. Paradiso. Chainmail: a configurable multimodal lining to enable sensate surfaces and interactive objects. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, TEI '10, pages 65–72, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-841-4. doi: <http://doi.acm.org/10.1145/1709886.1709899>.
- David Mizell. Using gravity to estimate accelerometer orientation. In *Proceedings of the Seventh IEEE International Symposium on Wearable Computers*, 2003. URL <http://csdl2.computer.org/comp/proceedings/iswc/2003/2034/00/20340252.pdf>.

- Aamer Mohamed, Ying Weng, Jianmin Jiang, and Stan Ipson. Face detection based neural networks using robust skin color segmentation. *Systems, Signals and Devices, 2008. IEEE SSD 2008. 5th International Multi-Conference on*, pages 1–5, July 2008. doi: 10.1109/SSD.2008.4632827.
- Bob Moog. Electronic music - its composition and performance. *Electronics World*, 1967.
- Anja Naumann, Jörn Hurtienne, Johann Habakuk Israel, Carsten Mohs, Martin Christof Kindsmüller, Herbert A. Meyer, Steffi Hußlein, and The IUUI Research Group. Intuitive use of user interfaces: Defining a vague concept. In *Engineering Psychology and Cognitive Ergonomics*, pages 128–136. Springer Berlin / Heidelberg, 2007.
- Dan Newton and Mark T. Marshall. Examining how musicians create augmented musical instruments. In *NIME '11: Proceedings of the 11th international conference on New interfaces for musical expression*, 2011.
- Kia Ng and Paolo Nesi. i-maestro framework and interactive multimedia tools for technology-enhanced learning and teaching for music. *Automated solutions for Cross Media Content and Multi-channel Distribution, 2008. AXMEDIS '08. International Conference on*, pages 266–269, Nov. 2008. doi: 10.1109/AXMEDIS.2008.41.
- K. Oka, Y. Sato, and H. Koike. Real-time fingertip tracking and gesture recognition. *IEEE Computer Graphics and Applications*, 22(6):64–71, 2002.
- Jaime. E. Oliver. The mano controller: A video based hand tracking system. In *International Computer Music Conference*, 2010.
- OED Online. *OED Online*. Oxford University Press, March 2010.
- Trond Are Øritsland and Jacob Buur. Taking the best from a company history - designing with interaction styles. In *DIS '00: Proceedings of the 3rd conference on Designing interactive systems*, pages 27–38, New York, NY, USA, 2000. ACM. ISBN 1-58113-219-0. doi: <http://doi.acm.org.ezproxy.sussex.ac.uk/10.1145/347642.347658>.
- Garth Paine. Interactivity, where to from here? *Org. Sound*, 7(3):295–304, 2002. ISSN 1355-7718. doi: <http://dx.doi.org/10.1017/S1355771802003096>.
- C Palacio-Quintin and M Zadel. Interactive composition and improvisation on the hyper-flute. In *International Computer Music Conference*, 2008.
- Daniel Perret. Roots of musicality: on neuro-musical thresholds and new evidence for bridges between musical expression and 'inner growth'. *Music Education Research*, 6(3):327 – 342, November 2004.
- Trevor J. Pinch and Karin Bijsterveld. Should one applaud?: Breaches and boundaries in the reception of new technology in music. *Technology and Culture*, 44(3):536–559, 2003.
- Cornelius Poepel. On interface expressivity: a player-based study. In *NIME '05: Proceedings of the 2005 conference on New interfaces for musical expression*, pages 228–231, Singapore, Singapore, 2004. National University of Singapore.
- Ivan Poupyrev, Michael J. Lyons, Sidney Fels, and Tina Blaine. New interfaces for musical expression. In *CHI '01 extended abstracts on human factors in computing systems*, 2001.
- Jennifer Preece, Yvonne Rogers, and Helen Sharp. *Interaction Design: Beyond Human Computer Interaction*. Wiley, 2002.
- P. Premaratne and Q. Nguyen. Consumer electronics control system based on hand gesture moment invariants. *Computer Vision, IET*, 1(1):35–41, March 2007. ISSN 1751-9640. doi: 10.1049/iet-cvi:20060198.
- Jeff Pressing. Cybernetic issues in interactive performance systems. *Computer Music Journal*, 14(1):12–25, 1990.
- Miller Puckette. Max at seventeen. *Comput. Music J.*, 26(4):31–43, 2002. ISSN 0148-9267. doi: <http://dx.doi.org/10.1162/014892602320991356>.

- Miller S. Puckette. Pure data: Another integrated computer music environment. In *in Proceedings, International Computer Music Conference*, pages 37–41, 1996.
- Wei Qiao. Echo-state-network-based real-time wind speed estimation for wind power generation. In *IJCNN'09: Proceedings of the 2009 international joint conference on Neural Networks*, pages 1505–1512, Piscataway, NJ, USA, 2009. IEEE Press. ISBN 978-1-4244-3549-4.
- Jeff Raskin. Intuitive equals familiar. *Communications of the ACM*, 37:9:17, 1994. URL <http://www.asktog.com/papers/raskintuit.html>.
- Michael Reed. Prototyping digital clay as an active material. In *TEI'09: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, pages 339–342, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-493-5. doi: <http://doi.acm.org/10.1145/1517664.1517733>.
- Mark Richardson. Autechre. <http://pitchfork.com/features/interviews/6788-autechre/>, 2008.
- Richard Riding. The nature and effects of cognitive style. In Robert J. Steinberg and Li-Fang Zhang, editors, *Perspectives on Thinking, Learning and Cognitive Styles*. Lawrence Erlbaum Associates, 2001.
- Richard Riding and Stephen Rayner. *Cognitive Styles and Learning Strategies: Understanding Style Differences in Learning and Behaviour*. David Fulton Publishers, 1998.
- Curtis Roads. A conversation with james a. moorer. *Computer Music Journal*, 6(4):10–21, 1982.
- Tara Rodgers. *Pink Noises*. Duke University Press, 2010.
- Joseph Butch Rovin, Marcelo M. Wanderley, Shlomo Dubnov, and Philippe Depalle. Instrumental gestural mapping strategies as expressivity determinants in computer music performance. In *Kansei - The Technology of Emotion Workshop*, 1997.
- Joel Ryan. Some remarks on musical instrument design at STEIM. *Contemporary Music Review*, 6(1): 3–17, 1991.
- Toshiki Sato, Haruko Mamiya, Hideki Koike, and Kentaro Fukuchi. Photoelastictouch: transparent rubbery tangible interface using an lcd and photoelasticity. In *UIST'09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pages 43–50, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-745-5. doi: <http://doi.acm.org/10.1145/1622176.1622185>.
- C Schwesig. What makes an interface feel organic? *Communications of the ACM*, 51(6), 2008.
- Allan Seago, Simon Holland, and Paul Mulholland. Timbre space as synthesis space: towards a navigation based approach to timbre specification. In *Spring Conference of the Institute of Acoustics*, 2008.
- Jennifer G. Sheridan and Nick Bryan-Kinns. Designing for performative tangible interaction. *International Journal of Arts and Technology*, 1(3/4):288–308, 2008.
- B. Shneiderman, G. Fischer, M. Czerwinski, M. Resnick, B. Myers, L. Candy, E. Edmonds, M. Eisenberg, E. Giaccardi, and T. Hewett. Creativity support tools: Report from a u.s. national science foundation sponsored workshop. *International Journal of Human-Computer Studies*, 20(2):61–77, 2006.
- R. T. Smith, B. H. Thomas, and W. Piekarski. Tech note: Digital foam. In *IEEE Symposium on 3D User Interfaces*, 2008a.
- Ross T. Smith, Bruce H. Thomas, and Wayne Piekarski. Digital foam interaction techniques for 3d modeling. In *VRST'08: Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pages 61–68, New York, NY, USA, 2008b. ACM. ISBN 978-1-59593-951-7. doi: <http://doi.acm.org/10.1145/1450579.1450592>.
- Sophy Smith. Compositional strategies of the hip-hop turntablist. *Organised Sound*, 5(2):75–79, 2000.
- Andrew Sorensen. Impromptu: An interactive programming environment for composition and performance. In *Proceedings of the Australasian Computer Music Conference*, pages 149–153, 2005.

- Andrew Sorensen and Andrew Brown. aa-cell in practice: An approach to musical live coding. In *Proceedings of the International Computer Music Conference*, 2007.
- Robert J. Steinberg. A capsule history of theory and research on styles. In Robert J. Steinberg and Li-Fang Zhang, editors, *Perspectives on Thinking, Learning and Cognitive Styles*. Lawrence Erlbaum Associates, 2001.
- D. Stowell, A. Robertson, N. Bryan-Kinns, and M. D. Plumbley. Evaluation of live human-computer music-making: quantitative and qualitative approaches. *International Journal of Human-Computer Studies*, 2009.
- Anselm Strauss. *Qualitative Analysis For Social Scientists*. Cambridge University Press, 1987.
- Anselm Strauss and Juliet Corbin. *Basics of Qualitative Research*. SAGE Publications, second edition, 1998.
- David Toop. Replicant: On dub. In *Audio Culture: Readings In Modern Music*. Continuum Books, 2002.
- Cathy P. Treadaway. Hand e-craft: an investigation into hand use in digital creative practice. In *C and C '09: Proceeding of the seventh ACM conference on Creativity and cognition*, pages 185–194, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-865-0. doi: <http://doi.acm.org/10.1145/1640233.1640263>.
- S. E. Trehub. The developmental origins of musicality. *Nature Neuroscience*, 6(7):669–673, 2003.
- Phil Turner. Towards an account of intuitiveness. *Behav. Inf. Technol.*, 27(6):475–482, 2008. ISSN 0144-929X. doi: <http://dx.doi.org/10.1080/01449290701292330>.
- Doug Van Nort and Marcelo Wanderley. Control strategies for navigation of complex sonic spaces. In *NIME '07: Proceedings of the 7th international conference on New interfaces for musical expression*, pages 379–382, New York, NY, USA, 2007. ACM. doi: <http://doi.acm.org/10.1145/1279740.1279824>.
- Michael E. Veal. *Dub: Soundscapes and Shattered Songs in Jamaican Reggae*. Wesleyan University Press, 2007.
- David Verstraeten. *Reservoir Computing: computation with dynamical systems*. PhD thesis, Ghent University, 2009.
- M. Waiswiz. The hands, a set of remote midi-controllers. In *Proceedings of the 1985 International Computer Music Conference*, pages 313–318. Computer Music Association, 1985.
- M. M. Wanderley and P. Depalle. Gestural control of sound synthesis. In *Proceedings of the IEEE*, volume 92, pages 632–644, 2004.
- Marcelo M. Wanderley and Marc Battier. *Trends in Gestural Control of Music*. Ircam - Centre Pompidou, 2000.
- Marcelo M. Wanderley, Bradley W. Vinesa, Neil Middletona, Cory McKay, and Wesley Hatch. The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, 34(1):97–113, 2005.
- Marcelo Mortensen Wanderley and Nicola Orio. Evaluation of input devices for musical expression: Borrowing tools from hci. *Comput. Music J.*, 26(3):62–76, 2002. ISSN 0148-9267. doi: <http://dx.doi.org/10.1162/014892602320582981>.
- Ge Wang and Perry R. Cook. ChucK: a concurrent, on-the-fly audio programming language. In *Proceedings of the International Computer Music Conference*, pages 219–226, 2003.
- Ge Wang and Perry R. Cook. On-the-fly programming: Using code as an expressive musical instrument. In *NIME '04: Proceedings of the 2004 conference on New interfaces for musical expression*, 2004.
- Robert Y. Wang and Jovan Popović. Real-time hand-tracking with a color glove. *ACM Trans. Graph.*, 28:63:1–63:8, July 2009. ISSN 0730-0301. doi: <http://doi.acm.org/10.1145/1531326.1531369>.

- Gil Weinberg and Seum-Lim Gan. The squeezables: Toward an expressive and interdependent multi-player musical instrument. *Computer Music Journal*, 25(2):37–45, 2001.
- David Wessel and Matthew Wright. Problems and prospects for intimate musical control of computers. In *Computer Music Journal*, pages 11–22, 2001.
- Frank R. Wilson. *The Hand*. Vintage Books, 1998.
- Todd Winkler. *Composing Interactive Music: Techniques and Ideas Using Max*. The MIT Press, 1998.
- Matthew Wright. Open sound control: an enabling technology for musical networking. *Organised Sound*, 10(3):193–200, 2005.
- Yaoyang Zheng and Kunio Shimada. Research on a haptic sensor made using mcf conductive rubber. *Journal of Physics: Condensed Matter*, 20(20):204148 (5pp), 2008. URL <http://stacks.iop.org/0953-8984/20/204148>.
- Hailing Zhou, Lijun Xie, and Xuliang Fang. Visual mouse: Sift detection and pca recognition. In *International Conference on Computational Intelligence and Security*, 2007.